frontiers | Research Topics

Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots

Edited by
Karl Friston, Georg Northoff, Tony J. Prescott,
Emily S. Cross and Sheri Marina Markose

Published in
Frontiers in Robotics and AI
Frontiers in Artificial Intelligence
Frontiers in Neurorobotics

ISSN 1664-8714
ISBN 978-2-8325-7159-0
DOI 10.3389/978-2-8325-7159-0

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots

Topic editors

Karl Friston — University College London, United Kingdom
Georg Northoff — University of Ottawa, Canada
Tony J. Prescott — The University of Sheffield, United Kingdom
Emily S. Cross — ETH Zürich, Switzerland
Sheri Marina Markose — University of Essex, United Kingdom

Citation

Friston, K., Northoff, G., Prescott, T. J., Cross, E. S., Markose, S. M., eds. (2026). *Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-7159-0

Table of contents

frontiers | Frontiers in Robotics and AI

TYPE Editorial
PUBLISHED 09 January 2026
DOI 10.3389/frobt.2025.1766766

OPEN ACCESS

EDITED AND REVIEWED BY
Chenguang Yang,
University of Liverpool, United Kingdom

*CORRESPONDENCE
Sheri Markose,
scher@essex.ac.uk

RECEIVED 13 December 2025
REVISED 13 December 2025
ACCEPTED 16 December 2025
PUBLISHED 09 January 2026

CITATION
Markose S, Prescott T, Northoff G, Cross E and Friston K (2026) Editorial: Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots.
Front. Robot. AI 12:1766766.
doi: 10.3389/frobt.2025.1766766

Editorial: Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots

Sheri Markose[1]*, Tony Prescott[2], Georg Northoff[3], Emily Cross[4] and Karl Friston[5]

[1]University of Essex, Colchester, United Kingdom, [2]The University of Sheffield, Sheffield, United Kingdom, [3]University of Ottawa, Ottawa, Canada, [4]Eidgenossische Technische Hochschule Zurich, Zürich, Switzerland, [5]University College London, London, United Kingdom

KEYWORDS

adaptive immune system, algorithmic takeover of biology, allostasis, artificial general intelligence (AGI), autonomy of goal setting, blockchain information processing, embodied self-referential genomic intelligence

Editorial on the Research Topic
Narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots

A team of multi-disciplinary editors, whose views are reflected in the themes underscored in this Research Topic, has come together to help take stock of the phenomenal success of narrow Statistical Artificial Intelligence (SAI) and to examine new perspectives on achieving Artificial General Intelligence (AGI). The tenets of SAI, which remain contingent on the domain of knowledge uploaded in digital form, is often contrasted with a long-held view that AGI must emulate the human brain, which marks an apogee as a prototype for general intelligence (Prescott, 2024). Many of the editors are of the view that the general scope of human intelligence arises from the necessity of maintaining the homeostasis of life itself (Friston, 2010; Friston, 2013; Prescott and Jimenez-Rodriguez, 2025), under ever-changing and often hostile circumstances. Further, they hold the view that complex life manifests embodied self-referential information processing (Northoff et al., 2006) with empathic mirror systems for prolific self-other interaction, and with selfhood and autonomy of goal setting intrinsically configured and committed to a hack-free agenda to mitigate what is inimical to life.

Against this background, we will briefly review the dozen papers published in this Research Topic involving a wide array of perspectives of 37 authors. These papers have to date garnered over 50,000 views and downloads.

It is useful to start with the review paper by Wu et al. where the scene is set for AI generations that have unfolded over the last 7 decades from AI 1.0 to AI 4.0. These developments have been driven by a triad of factors relating to algorithms and software; chip technology and computing power; access and storage of voluminous data in static and real time mode. AI 1.0 is ground zero, with algorithms aimed to fully direct outcomes mostly based on logic and rules based inference. This phase was accompanied by internet

technologies such as search engines, digital automation and data processing. The authors categorize AI 2.0 as encompassing Agentic AI, real-time online bots, and the advent of high-performance Graphics Processing Units (GPUs) and vast labeled datasets. This gave rise to deep learning and reinforcement learning, with convolutional and recurrent neural networks achieving breakthroughs in vision, language, and control. The third AI 3.0 generation marks the embedding of digital intelligence in an embodied physical agency of robots, where AI operates in material spaces as in autonomous vehicles and other utilitarian cobots.

The fourth generation AI 4.0 is set to coincide with aspirations of AGI with controversial notions of machine sentience, with self-capable, adaptive selection of goals and the wherewithal to evolve programs to achieve goals autonomously. Markose. points out that the lack of AI alignment (Bostrom, 2014; Russell, 2019) with human values and goals encountered in agents capable of setting their own goals is not unique to AI, but is a problem that lies at the foundations of civil society. Conflicting or adversarial goals of agents and their accompanying actions that are inimical to life must be kept in check for the survival of the human condition. On the other hand, providing AIs that are value-aligned with some awareness, and capacity for moral reasoning, could make them safer and better able to recognise and mitigate risks (Wallach, 2008).

The Markose. perspective on genomic intelligence—underpinned by the algorithmic takeover of biology within a uniquely encoded system—is that there are lessons to be learnt on the alignment problem from the evolution of general intelligence in complex life. The view here is that alignment to life and the design of selfhood has been solved in formal ways that can be explicated using Gödel logic, and with recent developments in cryptography with the blockchain. The principles involved here can be conjectured to maintain the immutability of original protein coding blocks against internal and external bio-digital adversaries within an evolvable and unbroken chain of life.

It has become popular to refer to self-improving code-based systems as Gödel machines in AGI frameworks, which are necessarily end-to-end self-assembly programs as in life (Schmidhuber, 2006; Zhang et al., 2025). Markose. suggests that this misses the point of Gödel logic, which is embedded in complex life first found in the adaptive immune system, AIS, of jawed fish 500 mya and latterly in the mirror neuron systems of primates. About 85% of expressed genes that can be identified as online self-assembly machines that create the morphology and phenotype of a multicellular organism can be viewed as its theorems. These are mapped offline in AIS 'Thymic Self' à la Self-Representation (Self-Rep) structures from the Recursive Function Theory of Gödel-Turing-Post. The purpose of this is to recursively identify non-self codes, especially of digital adversaries wielding the negation operator, which are potentially uncountable infinity. A corresponding open-ended capacity to detect changes to self-codes - known to be found only in the AIS and the human brain in a process of prolific predictive coding—is empowered by the Recombination Activation Gene operators. In a bold hypothesis, Markose (2022), Markose (2021) states that the Gödel Sentence is known to have little relevance in the real world, but is ubiquitous in complex life as a hashing algorithm (to adopt the language of blockchains) that enables embodied self-referential intelligence to detect any misalignment or negation of life's self-codes. This is accompanied by an arms race in novelty or surprises in a game with the viral/digital adversary, first identified by the game theorist Binmore (1987) in the archetype of Godel's Liar, to maintain the primacy of life codes. This self-regulation is achieved internally or by human external phenotypical interventions with human artifacts often in a structure of a perpetual Gödelian arms race.

This nicely takes us to other papers that investigate self-regulation and embodied intelligence within humans and AI systems. The research paper of Verchure et al. investigates the self-regulatory processes not through code-based smart controls, which can suffer misalignment by attacks by internal or external bio-malware as per Markose., but via the notion of allostasis, modelled by dynamical equations, whereby multiple physiological parameters are monitored and controlled "to maintain the stability of the integrated self rather than its parts". In particular, they consider how the mammalian brain conducts allostatic regulation of action, as an extension of the principle of homeostasis, using a predictive and adaptive multi-layered control architecture (see also Prescott and Jimenez-Rodriguez, 2025). They deploy an allocentric synthetic agent in a virtual environment and test the dynamical properties of the neural mass allostatic model with internal needs such as heat and hydration to be fulfilled in three scenarios. These relate to (1) open field rodent behavior, (2) where adaptation in navigation is needed, and (3) when criticality reset optimizes the interoceptive-driven decision-making process. They find, that though environmental stressors challenge the capacity to fulfil the agent's internal needs, the neural mass model with its self-regulatory dynamics achieves a robust balance in this regard.

The perspective paper by Caucke et al. explores how our understanding of the prolific capacity of social cognition in humans can help build the same capacities in robots. They review well-known theories on embodied self with self-knowledge - both from the interoceptive internal environment and the external environment, via the sensory motor cortex that undergirds physical situatedness. The use of self-knowledge as the basis of social cognition, empathy and action prediction of other similarly wired-up conspecifics and the strategic necessity of the Sally-Ann problem of false beliefs relating to perception of negation - are discussed. The authors are keen to emphasize that as human social cognition depends on some degree of individual autonomy, remote or externally controlled robots do not engage in social cognition. Likewise, they state that swarm robots that can self-organize along a well-defined and externally limited action set do not have autonomy in the choice of goals or actions. They touch on the fundamental problem of coordination and cooperation when robot behaviors are mutually predictable by robots themselves via good internal models of the other. This requires that the robots do not engage in unpredictable actions that are adversarial or disruptive of what is mutually predictable. While specific robots can have their autonomy limited in order to be cooperative, as indicated in the seminal work of Binmore (1987), digital adversaries cannot be eliminated in general and robots like humans must be capable of detecting Liars/adversaries and enter into arms races with them to preserve autonomy of self.

In Ryan. perspective paper, the embodied and ecological approach to intelligence, with the former based on the framework of the Learning Intelligent Decision Agent (LIDA), is used to understand novelty and improvisation in music. For this Ryan

draws on the Jeff Pressing model which entails the knowledge base comprised of cognitive units of objects and processes stored in long term memory of the musician. Processes bring about changes in features and objects and all of these aspects interact with one another in complex ways. Ryan favors the LIDA approach rather than AGI as LIDA's framework of cognitive cycles of learning, perception, and action engages high and low levels of cognition of an autonomous agent without a specific problem-solving goal typically associated with AGI. In future works Ryan aims to show a LIDA/Pressing robot for music improvisation compared with existing improvising AI machines.

Pontes-Filho et al. challenge the view that AGI should emulate human-level intelligence and instead argue that the starting point of AGI should be at a much lower level, which they call Neuroevolutionary AGI (NAGI) when learning occurs through sensory experience. They propose at a minimum - a body and a reactive environment - where evolutionary complexification can happen. From a randomly initialized spiking neural network, they posit that learning occurs with adaptive synapses which control binary signals (excitatory or inhibitory) that propagate through reconfigurable network topology. Hence, this method has been called Neuroevolution of augmented topologies (NEAT). This method, though comparable, does not follow gradient descent of deep neural networks for reinforcement learning tasks. NAGI is successfully tested on three tasks: food foraging, emulation of logic gates, and cart-pole balancing. This approach, while promising, begs the question of how an AI learning to do preassigned tasks can achieve AGI ambitions, which typically include autonomous choice of goals themselves.

Johansson. aims to advance AGI by developing what he calls Machine Psychology by harnessing operant learning psychology based on behavioral changes due to the consequences of actions, integrating it with the Non-Axiomatic Reasoning System (NARS). NARS has been built with sensorimotor reasoning at its core, enabling it to process sensory data in real-time and respond with appropriate motor actions. NARs is equipped to be an efficacious inference system with limited knowledge and resources, a condition that is often true for real-world scenarios. Combining the two is an apt example of learning by doing, though Johansson brings in the Skinnerian behavioral triad of stimulus, response, and reward and an additional *establishing operation* (EO), which can enhance or mitigate the stimulus and make the response more (less) likely. Well-known critiques of Skinner—such as that by Chomsky—exist on how language acquisition, for instance, requires more than operant conditioning. The tasks used to test out Machine Psychology, though successful, fall far short of the prowess expected by AGI. The clear advantage of NARS is that it can eschew large data sets, unlike traditional AI systems, as NARS operates effectively under conditions of insufficient knowledge and resources.

In robot intelligence Chen et al. like Johansson. propose the use of fast and frugal heuristic decision making, as in humans, conjectured to lead to more robust inference in real-time systems in which rapid decision making is essential. Bounded rationality solutions of Herbert Simon that rely on satisficing rather than optimization underpins the branch of learning called active perception in robots, which uses less data than onerous deep learning solutions. The authors use human decision makers to solve simulated treasure hunt problems in a virtual environment to derive efficacious decision rules as time and other pressures, such as impediments to visual perceptions (fog), are increased. The most efficacious human strategies discovered from human studies are then implemented on autonomous robots equipped with vision sensors. Results show robust performance of robots using the heuristic toolbox, when compared with known optimization algorithms that fail to complete the search for treasures under unanticipated adverse conditions.

In a second paper, Johansson. shows how Arbitrarily Applicable Relational Responding (AARR)—which has been considered to be a particularly human facility for flexible and contextual learning—can be captured by suitably designed AI systems. He aims to achieve this by combining Non-Axiomatic Reasoning System (NARS) used for learning under uncertainty with the behavioral psychology account of AARR, which enables NARS to derive symbolic relational knowledge directly from sensorimotor experiences. He shows how key properties of AARR (mutual entailment, combinatorial entailment, and transformation of stimulus functions) can emerge from NARS's inference rules and memory structures. The claim is that this can pave the way to AGI, though there is some considerable work that needs to be done to bring this to fruition.

In the final three papers that are reviewed here, applications of extant AI for specific tasks are considered, or some new enhancements have been incorporated to achieve more efficacious performance.

Zhai et al. use a multi-modal and multi-level approach for enhanced human-robot interactions. Multi-modal intelligence is a desired AGI characteristic, as a combination of visual, auditory, and language-enabled human intelligence gives enhanced experience and performance, and impairment in any of these modalities places an individual at a considerable disadvantage in life. However, in a robot setting, the architecture of multimodal intelligence is considerably more complicated when combining, say, computer vision for object identification with natural language processing for named entity identification. This is especially the case in settings like social media postings, where images and texts are short and prone to noise, making it harder to achieve feature selection and identify relevant information. The authors develop a multimodal named entity recognition (MNER) architecture in which the neural network can extract useful visual information for enhancing semantic understanding and subsequently improve entity identification. Twitter data sets with pictures and text are used for experiments and to test out their MNER model. The enhanced performance of their MNER—when compared to other multi-modal models—comes at the price of slower operations.

Babushkin et al. investigate how handwriting, which is the outcome of multimodal inputs in humans, can be evaluated by an AI such as a temporal convolutional neural network (TCN). The use of AI is seen to overcome biases that humans have in their assessment of handwriting; there are educational, forensic, and technological contexts where AI can provide a more efficient and accurate service. For their AI experiment, the handwritten documents of an identical text—designed to include all possible orthographic combinations of Arabic characters—were done by 50 human subjects, and three experts were used to categorize the produced text into different legibility scores. The TCN is trained to classify the documents into different legibility categories. The results show that while the TCN model trained on stylus kinematics features

demonstrates relatively high accuracy (around 76%), the addition of hand kinematics features significantly increases the model accuracy by approximately 10%.

In the final paper Binzagr and Abulfaraj. aim to improve traditional machine learning methods for diagnosing Alzheimer's disease (AD) from MRI scans, claiming that CNN architectures have problems with detecting AD due to overfitting. In addition to the CNN, the authors incorporate two other components into their new framework. These include a generalized self-attention (GSA) score, which gives a global assessment of interdependence across spatial and channel dimensions while filtering out irrelevant details, and an extreme learning machine (ELM) classifier, employed to categorize AD. Note, the GSA blocks are placed on an InceptionV3 network, which is a directed acyclic graph (DAG) network that has 316 layers and 350 links that include 94 as convolutional layers. In-depth experiments on two benchmark datasets demonstrate that the proposed InGSA achieves superior performance compared to the state-of-the-art techniques.

This Research Topic provides a valuable snapshot of where thinking about AGI is in the early-mid 2020s. Some crosscutting themes in the research and review articles laid out above are agency, autonomy, and autopoiesis, read in their broadest terms. Indeed, 'codeopoiesis' or how code-based genomic intelligence achieves self-organization (Markose, 2022) as in blockchains is underscored to reflect new developments in autonomous AI.

Author contributions

SM: Conceptualization, Writing – original draft. TP: Writing – original draft, Writing – review and editing. GN: Writing – review and editing. EC: Writing – review and editing. KF: Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors SM, TP, GN, EC, and KF declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Bostrom, N. (2014). *Superintelligence: paths, dangers, strategies*. New York, NY: OUP.

Binmore, K. (1987). Modeling rational players: part I. *Economics and Philosophy* 3 (2) 179–214.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi:10.1038/nrn2787

Friston, K. (2013). Life as we know it. *J. R. Soc. Interface* 10, 20130475. doi:10.1098/rsif.2013.0475

Markose, S. M. (2021). Genomic Intelligence as Über Bio-Cybersecurity: the Gödel Sentence in Immuno-Cognitive Systems. *Entropy* 23, 405. doi:10.3390/e23040405

Markose, S. M. (2022). Complexification of eukaryote phenotype: adaptive immuno-cognitive systems as unique gödelian blockchain distributed ledger. *Biosystems* 220, 104718. doi:10.1016/j.biosystems.2022.104718

Northoff, G., Heinzel, A., Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage* 31 (1), 440–457. doi:10.1016/j.neuroimage.2005.12.002

Prescott, T. J. (2024). *The psychology of artificial intelligence*. London: Routledge.

Prescott, T. J., and Jimenez-Rodriguez, A. (2025). "Understanding the layered brain architecture for motivation: dynamical systems, computational neuroscience, and robotic approaches," *Psychol. Learn. Motivation*, 62. 45–96. doi:10.1016/bs.plm.2025.03.005

Russell, S. (2019). *Human compatible: artificial intelligence and the problem of control*. Penguin Random House.

Schmidhuber, J. (2006). Gödel machines: self-referential universal problem solvers making provably optimal self-improvements. *arXivcs/0309048*. doi:10.48550/arXiv.cs/0309048

Wallach, W. (2008). *Moral machines: teaching robots right from wrong*. Oxford: OUP.

Zhang, J., Hu, S., Lu, C., Lange, R., and Clune, J. (2025). Darwin Gödel machine: open-ended evolution of self-improving agents. *arXiv:2505.22954* 2. doi:10.70777/si.v2i3.15063

frontiers | Frontiers in Robotics and AI

TYPE Original Research
PUBLISHED 14 October 2022
DOI 10.3389/frobt.2022.1007547

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Stefania Costantini,
University of L'Aquila, Italy
Neil Vaughan,
University of Exeter, United Kingdom

*CORRESPONDENCE
Sidney Pontes-Filho,
sidneyp@oslomet.no

SPECIALTY SECTION
This article was submitted to Computational Intelligence in Robotics, a section of the journal Frontiers in Robotics and AI

RECEIVED 30 July 2022
ACCEPTED 03 October 2022
PUBLISHED 14 October 2022

CITATION
Pontes-Filho S, Olsen K, Yazidi A, Riegler MA, Halvorsen P and Nichele S (2022), Towards the Neuroevolution of Low-level artificial general intelligence. *Front. Robot. AI* 9:1007547.
doi: 10.3389/frobt.2022.1007547

Towards the Neuroevolution of Low-level artificial general intelligence

Sidney Pontes-Filho[1,2]*, Kristoffer Olsen[3], Anis Yazidi[1,4,5], Michael A. Riegler[6,7], Pål Halvorsen[1,6] and Stefano Nichele[1,4,5,6,8]

[1]Department of Computer Science, Oslo Metropolitan University, Oslo, Norway, [2]Department of Computer Science, Norwegian University of Science and Technology, Trondheim, Norway, [3]Department of Informatics, University of Oslo, Oslo, Norway, [4]AI Lab—OsloMet Artificial Intelligence Lab, Oslo, Norway, [5]NordSTAR—Nordic Center for Sustainable and Trustworthy AI Research, Oslo, Norway, [6]Department of Holistic Systems, Simula Metropolitan Centre for Digital Engineering, Oslo, Norway, [7]Department of Computer Science, UiT the Arctic University of Norway, Tromsø, Norway, [8]Department of Computer Science and Communication, Østfold University College, Halden, Norway

In this work, we argue that the search for Artificial General Intelligence should start from a much lower level than human-level intelligence. The circumstances of intelligent behavior in nature resulted from an organism interacting with its surrounding environment, which could change over time and exert pressure on the organism to allow for learning of new behaviors or environment models. Our hypothesis is that learning occurs through interpreting sensory feedback when an agent acts in an environment. For that to happen, a body and a reactive environment are needed. We evaluate a method to evolve a biologically-inspired artificial neural network that learns from environment reactions named Neuroevolution of Artificial General Intelligence, a framework for low-level artificial general intelligence. This method allows the evolutionary complexification of a randomly-initialized spiking neural network with adaptive synapses, which controls agents instantiated in mutable environments. Such a configuration allows us to benchmark the adaptivity and generality of the controllers. The chosen tasks in the mutable environments are food foraging, emulation of logic gates, and cart-pole balancing. The three tasks are successfully solved with rather small network topologies and therefore it opens up the possibility of experimenting with more complex tasks and scenarios where curriculum learning is beneficial.

KEYWORDS

neuroevolution, artificial general intelligence, spiking neural network, spike-timing-dependent plasticity, Hebbian learning, weight agnostic neural network, meta-learning

1 Introduction

Artificial General Intelligence (AGI) or strong Artificial Intelligence (AI) is commonly discussed among AI researchers. It is often defined as human-level AI. However, the generality of an AI does not need to be considered at such a level of complexity. Even an artificial neural network that performs lots of different tasks as a collection of specialized

or weak AI (Reed et al., 2022) may not provide the level of generality observed in simple biological systems. In fact, our current artificial intelligent systems cannot emulate the adaptability to unknown conditions and learning capabilities of an animal with a simple nervous system, such as a worm (Ardiel and Rankin, 2010; Randi and Leifer, 2020). An alternative approach is to start the quest for the generality of AI from the simplest tasks that animals can do, but machines cannot, like behaving intelligently even in new environments (Crosby et al., 2019), i.e., out-of-distribution generalization (Shen et al., 2021). Moreover, AGI systems should be tested in tasks that require self-learning on the fly from sensory feedback, as it is often done in meta-learning and continual learning (Najarro and Risi, 2020; Zohora et al., 2021).

We argue that a radical paradigm change is needed in order to reach general intelligence (Lake et al., 2017; Crosby et al., 2019). Our hypothesis is that such a new paradigm requires learning systems with self-organizing properties, as discussed by Risi, (2021). In this work, our goal is to achieve the learning capabilities of a primitive brain. Therefore, we aim at a low-level AGI, i.e., a system that can learn a map function through sensory experience. Interpreting and understanding sensory inputs are achieved through evolution, particularly supervised evolution (Zador, 2019) of agents interacting with their environment.

The brain is the organ that interprets the encoded signals from our sensory organs, thanks to the ability to distinguish between positive and negative sensory experiences depending on what is considered to be good or harmful, e.g., pleasure and pain. The experiences of pleasure and pain serve as reward and penalty mechanisms that may affect our behavior by conditioning associative positive and negative cues with specific memories.

In this work, we evaluate the Neuroevolution of Artificial General Intelligence (NAGI) framework (Pontes-Filho and Nichele, 2019). NAGI is a low-level biologically-inspired AGI framework. NAGI consists of an evolvable spiking neural network with adaptive synapses and randomly-initialized weights. The network is evolved by an extension of the method NeuroEvolution of Augmenting Topologies (NEAT) (Stanley and Miikkulainen, 2002). The source code of NAGI is available at https://github.com/SocratesNFR/neat-nagi-python.

The evolved spiking neural network controls an agent placed in a mutable environment. Its chances of reproduction are proportional to how long it can survive in an environment that is constantly changing, sometimes abruptly. Evolution optimizes how the neurons are connected in the network, their type of neurotransmitters (excitatory or inhibitory), their susceptibility to background electrical current noise (analogous to bias), and their neuroplasticity. With such degrees of freedom in the optimization process, we attempt to approximately recapitulate the evolutionary process of the simplest brains. The mutable environment and random weight initialization propitiate a benchmark for generality and adaptivity of the agent.

We test NAGI in three mutable environments. The first one is a simple food foraging task, in which the agent has one photoreceptor (or light intensity sensor) used to identify food. The food type (color) is either black or white. Food can be edible or poisonous and this feature changes over time. The agent can also taste the food as its sensory feedback for good and bad actions. The second environment is a logic gate task. The spiking neural network needs to emulate different logic gates in series where the only reward and penalty sensory signals are the supporting mechanisms to identify the correct output. The third environment is a cart-pole balancing task. In this environment, the goal of the agent is to control the forces applied to the cart in order to maintain the pole above itself upright. The mutable component of this environment is the pole length, which changes during the lifetime of the agent. Because this environment has sensory feedback for the agent's actions, there is no need to add reward and penalty sensory signals.

The article is organized as follows: Section 2 explains the theoretical basis for understanding NAGI. Section 3 discusses the related work to our approach. Section 4 describes the details of the method and experiments. Section 5 presents the experimental results. Section 6 concludes the article including a discussion of the results and plans for future work.

2 Background

The components of the NAGI framework are inspired by the overlapping research fields of artificial life (Langton, 2019), evolutionary robotics (Doncieux et al., 2015), and computational neuroscience (Trappenberg, 2009). In particular, the controller for the agents is a Spiking Neural Network (SNN) (Izhikevich, 2003), which is a more biologically-plausible artificial neural network. The neurons in an SNN communicate through spikes, i.e., binary values in time series. Therefore, an SNN adds a temporal dimension to binary data. A neuron propagates such data depending on whether its membrane potential crossed a threshold value or not. If the threshold is crossed, the neuron propagates a signal represented as neurotransmitters to its connected neurons; otherwise, the action potential is not propagated. When neurotransmitters are released by a neuron, they can be of two types: excitatory, which increases the membrane potential and the likelihood of producing an action potential; or inhibitory, which has the opposite effect by decreasing the membrane potential. Efficient optimization of an SNN cannot happen through gradient descent as spike trains are not differentiable (Tavanaei et al., 2019). Instead, spiking neurons have biologically inspired local learning rules, such as Hebbian learning and Spike-Timing-Dependent Plasticity (STDP) (Hebb, 1949; Li et al., 2014). Those neuroplasticity rules are unsupervised, and their functionality in the brain is still not fully understood. However, it is inferred that the supervision comes

from a certain network configuration acquired through evolution. Therefore, in this work, we use a modification of NeuroEvolution of Augmenting Topologies (NEAT) (Stanley and Miikkulainen, 2002). NEAT uses a Genetic Algorithm (GA) (Holland, 1992) to optimize the weights and the topology of a growing neural network that is initialized with a minimal and functional size. NEAT is typically used to search for a network configuration that improves a fitness score while maintaining population diversity (speciation) and avoiding loss of genes during crossover (historical marking). For an accessible and extensive explanation of NEAT, please refer to Ref. (Welleck, 2019).

A distinction from NEAT is that the weights in the NAGI framework are randomly initialized, and they change (adapt) after deployment. The adaptation is coordinated by a realistic Hebbian learning rule, i.e., STDP. This neuroplasticity adjusts the synaptic strength of a neuron's dendrites (i.e., input connections) when it fires an action potential (or spike) that goes through its axon (i.e., output connection). The weights are modified according to the difference in time between incoming spikes and the generated action potential. More detailed information about SNN and STDP is available in Ref. (Camuñas-Mesa et al., 2019).

The body and brain interaction (sensors and actuators vs. controller) is often described as "chicken and egg" problem (Funes and Pollack, 1998). The natural evolution of body and brain happens together with the evolution of the environment. They evolve in cooperation and response to each other (Mautner and Belew, 2000). The application of supervised evolution of agents interacting with the environment is defined as embodied evolution (Watson et al., 1999). As such, an agent needs a body to learn from the reaction of its environment. We hypothesize that low-level general intelligence in nature emerged through the evolution of a sensory feedback learning method.

3 Related work

Neuroevolution with adaptive synapses was introduced in 2003 by Stanley et al. (2003). Such a method is a version of NEAT where the synaptic strength of the connections changes with Hebbian local learning rules. In their work, they used a food foraging task where an agent moves around a field surrounded by edible and poisonous food. The type of food did not change over time, but it was initialized differently at every new run. The agents needed to try the food first before identifying it. Therefore, the agents possess reward and penalty sensory signals as in NAGI. This method is rather similar to ours. However, NAGI is more biologically plausible, weight agnostic, and is tested in a mutable environment. Risi and Stanley, (2010) proposed an extended version by replacing the direct encoding of the network in NEAT with an indirect encoding.

Additional related methods are described in Refs. (Gaier and Ha, 2019) and (Najarro and Risi, 2020) where randomly-initialized artificial neural networks are used. The work of Gaier and Ha, (2019) uses a version of NEAT where each neuron can have one activation function out of several types. While in the method of Najarro and Risi, (2020), the network topology is fixed and each connection evolves to optimize the parameters of its Hebbian learning rule.

In a recent review on neuroevolution (Stanley et al., 2019), NEAT and its extensions are comparable to deep neural networks trained with gradient-based methods for reinforcement learning tasks. Such methods allow evolving artificial neural networks with indirect encoding for scalability, novelty search for diversity, meta-learning for learning how to learn, and architecture search for deep learning models. Moreover, neuroevolution is described as a key factor for reaching AGI, particularly in relation to meta-learning and open-ended evolution. Meta-learning encompasses the training of a model with certain datasets and testing it with others. The goal of the model is therefore to learn any given dataset by itself from experience (Thrun and Pratt, 1998). Open-ended evolution is the ability to endlessly generate a variety of solutions of increasing complexity (Taylor, 2019). In NAGI, meta-learning is an implicit target in the mutable environments and is implemented as neuroplasticity in the spiking neural network.

In 2020, Nadji-Tehrani and Eslami, (2020) introduced the framework for evolutionary artificial general intelligence (FEAGI). This method uses an indirect encoding technique for a spiking neural network that resembles the growth of the biological brain, which is called "neuroembryogenesis." As a proof of concept, FEAGI demonstrates successful handwritten digits classification by learning through association and being able to recall digits from different image samples in real-time.

4 Neuroevolution of Artificial General Intelligence

The NAGI framework aims at providing a simplified model of the initial stages of the evolution of biological general intelligence (Pontes-Filho and Nichele, 2019). The evolving agents in NAGI consist of randomly-initialized spiking neural networks. Thus, a genome in NAGI does not require the definition of synaptic weights of the connections between neurons, as it is done in NEAT. Therefore, the synaptic weights in the genome are replaced by an STDP rule and its parameters for each neuron. Since biological neurons may provide one of the two main neurotransmitters, NAGI's genome defines such a feature in the neurons' genes. As such, a neuron can be either excitatory or inhibitory. To imitate the function of bias in artificial neural networks, neurons may be also susceptible to a "background electrical current noise."

The environment changes during the lifetime of the agent. This forces the agent to learn new environmental conditions. Therefore, the agent is encouraged to generalize and learn how to learn. The aforementioned random initialization and mutable environment aim at benchmarking the basic properties needed for low-level AGI.

4.1 Spiking neural network

The spiking neural network has a fixed number of input and output neurons depending on the task to be solved. The neuroevolution process defines the number of hidden neurons that will be available. Hidden neurons can be either excitatory or inhibitory, while input and output neurons are always excitatory. Self-loops and cycles are permitted while duplicate connections between two neurons in the same direction are prohibited. The SNN is stimulated from the input neurons, as such units are spike generators. The spikes are uniformly generated in an assigned frequency or firing rate.

As a spiking neuron model, we use a simplification of the leaky integrate-and-fire model (Liu and Wang, 2001). A neuron's membrane potential v is increased directly by its inputs and decays over time by a factor λ_{decay}. We can then express the change in membrane potential Δv with regards to a time step Δt by

$$\Delta v(\Delta t) = \sum_{i=1}^{n} w_i x_i - \Delta t \lambda_{decay} v, \tag{1}$$

where x_i is the input value 0 (no spike) or 1 (spike) from the presynaptic neuron i, the dendrite for this connection has the synaptic strength defined as w_i, and n is the total number of presynaptic neurons that the dendrites are connecting. If the membrane potential v is greater than the membrane threshold v_{th}, a spike is released and the membrane potential returns to the resting membrane potential v_{rest}, which is 0. The time step Δt we use in the experiments is 0.1 ms, and decay factor λ_{decay} is $0.01\Delta t$. An action performed by the SNN is calculated by the number of spikes in a time window. Such an actuator time window covers 250 ms or 2,500 time steps. In NAGI, the weights of the SNN are randomly initialized with a normal distribution. The mean is equal to 1 and the standard deviation is equal to 0.2. The weights are always positive. As mentioned, the excitation and inhibition of a neuron are defined by the neurotransmitter of the presynaptic neuron.

4.1.1 Homeostasis

Biological neurons have a plasticity mechanism that maintains a steady equilibrium of the firing rate, which is called homeostasis (Betts et al., 2013; Kulik et al., 2019). In our method, the spiking neurons can have non-homogeneous inputs, which could lead to very different firing rates. It is desirable that all neurons have approximately equal firing rates (Diehl and Cook, 2015). In order to homogenize the firing rates of the neurons in a network, the membrane threshold v_{th} is given by

$$v_{th} = \min\left(v_{th}^{*} + \Theta, \sum_{i=1}^{n} w_i\right), \tag{2}$$

where v_{th}^{*} is the "resting" membrane threshold equals to 1; and Θ starts with value 0, increases 0.2 every time a neuron fires, and decays exponentially with a rate of $0.01\Delta t$. Each neuron has an individual Θ. Therefore, a neuron firing more often will get a larger membrane threshold and consequently a lower firing rate. To compensate for a neuron with weak incoming weights, which causes a low firing rate; we instead use the sum of the incoming weights as the threshold.

4.1.2 Spike-Timing-Dependent Plasticity

The adjustment of the weights of the connections entering into a neuron happens on every input and output spike to and from a neuron. This is performed by STDP. It is done by keeping track of the time elapsed since the last output spike and each input spike from incoming connections within a time frame. Such a time frame is called the STDP time window and is set to be ±40 ms. The difference between presynaptic and postsynaptic spikes, or the relative timing between them, denoted by Δt_r is given by

$$\Delta t_r\,(t_{out}, t_{in}) = t_{out} - t_{in}, \tag{3}$$

where t_{out} is the timing of the output spike and t_{in} is the timing of the input spike.

The synaptic weight change Δw is calculated in accordance with one of the four Hebbian learning rules. The functions for each of the four learning rules are given by

$$\Delta w(\Delta t_r) = \begin{cases} A_+ e^{\frac{-\Delta t_r}{\tau_+}} & \Delta t_r > 0, \\ -A_- e^{\frac{\Delta t_r}{\tau_-}} & \Delta t_r < 0, \\ 0 & \Delta t_r = 0; \end{cases} \quad \text{Asymmetric Hebbian} \tag{4}$$

$$\Delta w(\Delta t_r) = \begin{cases} -A_+ e^{\frac{-\Delta t_r}{\tau_+}} & \Delta t_r > 0, \\ A_- e^{\frac{\Delta t_r}{\tau_-}} & \Delta t_r < 0, \\ 0 & \Delta t_r = 0; \end{cases} \quad \text{Asymmetric Anti} - \text{Hebbian} \tag{5}$$

$$\Delta w(\Delta t_r) = \begin{cases} A_+ g(\Delta t_r) & g(\Delta t_r) > 0, \\ A_- g(\Delta t_r) & g(\Delta t_r) < 0, \\ 0 & g(\Delta t_r) = 0; \end{cases} \quad \text{Symmetric Hebbian} \tag{6}$$

$$\Delta w(\Delta t_r) = \begin{cases} -A_+ g(\Delta t_r) & g(\Delta t_r) > 0, \\ -A_- g(\Delta t_r) & g(\Delta t_r) < 0, \\ 0 & g(\Delta t_r) = 0; \end{cases} \quad \text{Symmetric Anti} - \text{Hebbian} \tag{7}$$

where $g\,(\Delta t_r)$ is a Difference of Gaussian function given by

$$g(\Delta t_r) = \frac{1}{\sigma_+\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\Delta t_r}{\sigma_+}\right)^2} - \frac{1}{\sigma_-\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\Delta t_r}{\sigma_-}\right)^2}, \tag{8}$$

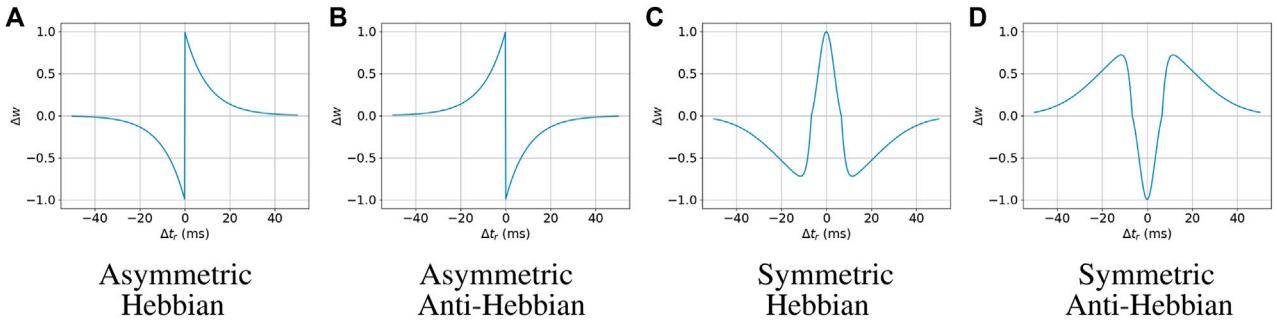


FIGURE 1
Spike-timing-dependent plasticity rules.

A_+ and A_- are the parameters that affect the height of the curve, τ_+ and τ_+ are the parameters that affect the width or steepness of the curve of the Asymmetric Hebbian functions, and σ_+ and σ_- are the standard deviations for the Gaussian functions used in the Symmetric Hebbian functions. It is also required that $\sigma_- > \sigma_+$. We experimentally found fitting ranges for each of these parameters, which are $A_+ = [0.1, 1.0]$, $A_- = [0.1, 1.0]$, $\tau_+ = [1.0, 10.0]$, and $\tau_- = [1.0, 10.0]$ for the asymmetric STDP functions; and $A_+ = [1.0, 10.6]$, $A_- = [1.0, 44.0]$, $\sigma_+ = [3.5, 10.0]$, and $\sigma_- = [13.5, 20.0]$ for the symmetric ones. The STDP curves with the maximum value of those parameters are illustrated in Figure 1.

Weights can take values in a range $[w_{\min}, w_{\max}]$, and every neuron has a weight budget w_{budget} it must follow. What this means is that if the sum of a neuron's incoming weights exceed w_{budget} after initialization or STDP has been applied, they are normalized to w_{budget}, given by

$$\text{if } \sum_{i=1}^{n} w_i > w_{budget}, \text{then } w_i = \frac{w_i w_{budget}}{\sum_{i=1}^{n} w_i}. \quad (9)$$

The parameters used during our experiments are $w_{\min} = 0$, $w_{\max} = 1$, and $w_{budget} = 5$. In case of a SNN without homeostasis, if a connection i has $w_i = w_{\max}$, then $w_i = v_{th}$. Therefore, an action potential coming from i will always produce a spike. This is the reason why $w_{\max} = v_{th}$.

4.2 Genome

The genome in NAGI is rather similar to the one in NEAT. Its node genes have three types: input, hidden, and output. Depending on the type of the node gene, there is a different collection of *loci*[1]. The input node is a spike generator and

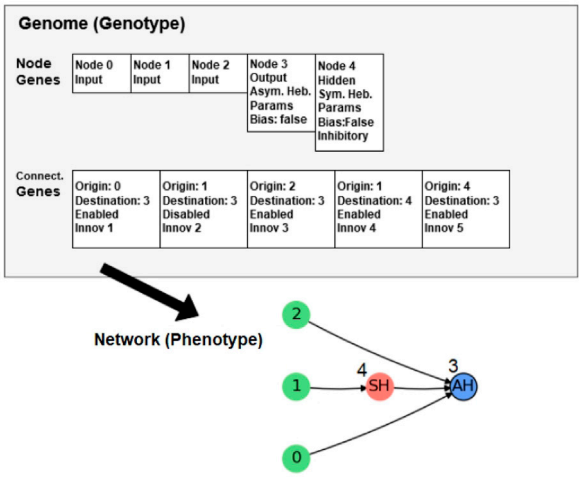


FIGURE 2
Genotype and phenotype in NAGI. Image taken from Ref. (Olsen, 2020).

provides excitation to the neurons it is connected to. The gene of an input node is the same as in NEAT. The hidden and output nodes represent adaptable and mutable spiking neurons. They have three additional loci: the type of the learning rule, the set of the learning rule parameters, and a bias. The connection gene in NAGI has no weight locus as in NEAT. The reason for its removal is that the weights of the SNN are defined by a normal distribution.

The learning rule is one of the four STDPs. The set of learning rule parameters consists of four parameters that adjust the intensity of the weight change. They are different for symmetric and asymmetric learning rules. The symmetric parameters are $\{A_+, A_-, \sigma_+, \sigma_-\}$ and the asymmetric parameters are $\{A_+, A_-, \tau_+, \tau_-\}$. The bias is a Boolean value that determines if the neuron has a constant input of 0.001 being added to Δv, which is analogous to the background noise of the neuron.

1 In the terminology of genetic algorithms, a value within a gene is also called a *locus* (plural *loci*).

The hidden node genes have a unique locus, which is a Boolean value that determines whether it represents an inhibitory or excitatory neuron. This locus is not included in the output node genes because they are always excitatory. As a result of combining all the descriptions of the genome in NAGI, the genotype and the phenotype are illustrated in Figure 2.

The initialization of the additional loci in the node genes can be conditional and non-uniform. The initialization of the neurotransmitter type of a neuron follows a similar proportion of excitatory and inhibitory neurons in the brain (Sukenik et al., 2021). The probability of a neuron being added as excitatory is 70%. The probability of having a bias is 20%. Depending on the neurotransmitter, excitatory neurons have a 70% chance of initializing with Hebbian plasticity, and inhibitory neurons have the same chance but for anti-Hebbian plasticity. The learning rule parameters are initialized by sampling from a uniform distribution within the STDP parameter ranges.

The mutations of the additional loci happen in 10% of chance to switch the neurotransmitter type, bias, learning rule, and learning rule parameters. Those parameters have 2% chance of a fully re-initialization. When the parameters are assigned to be mutated, a random value sampled from a normal distribution with $\mu = 0$ and $\sigma^2 = m(p)$ is added to the parameter p. The equation of $m(p)$ is

$$m(p) = 0.2\,(p_{max} - p_{min}), \tag{10}$$

where $p_{\max}$ and $p_{\min}$ are the maximum and minimum values the parameter can have, given by the STDP parameter ranges. During the neuroevolution, 10% of the genotypes with the best fitness scores will be passed to the next generation unchanged, i.e., elitism.

4.3 Mutable environments

The benchmark tasks for NAGI are meant to evaluate the agent's ability to generalize and self-adapt. Therefore, they consist of environments that change during the lifetime of the agent. Two types of tasks are provided, binary classification (two tasks of this kind are provided) and control (one task of this kind is provided). The first type (binary classification) is the simplest one, however, it provides the most abrupt changes in the environment. The binary classification tasks are food foraging with one input, and logic gates with two inputs. The control task in a simulated physical environment is the cart-pole balancing from OpenAI Gym (Brockman et al., 2016). The changes are less abrupt in this last task as they consist in modifying the pole size. The fitness scores are calculated using the number of time steps t that the agent survived in these environments, normalized to the range [0, 1] using the maximum possible lifetime $L_{\max}$ and minimum possible lifetime $L_{\min}$. Therefore, the fitness function f is given by

$$f(t) = \frac{t - L_{min}}{L_{max} - L_{min}}. \tag{11}$$

In the binary classification tasks, the agents have an initial amount of health points that is reduced every time step as continuous damage. If a correct action is chosen, the health point amount is reduced by d_c health point. Otherwise, it is reduced by d_i. The input sample is given to the agent for 1 s or 10,000 time steps, then it is changed to a new one. The mutation of the environment condition happens when the agent has seen four samples. The order of the input samples and the environment conditions is fixed and cyclic.

We noticed that the number of spikes within the actuator time window can be the same for the output neurons and therefore allowing for a tie in many cases. Our solution to avoid spiking neural networks with this behavior is to include a "confidence" factor in the fitness score calculation. Therefore, the higher the difference between the spike count, the more confident the action is. If the action is correct and highly confident, the damage is d_c or closer. If the action is incorrect but highly confident, the damage is d_i or closer. The lack of confidence would make the damage lie between the values d_c and d_i. The spike count for the correct action s_c and incorrect one s_i are used to calculate the participation of the spikes for deciding the correct action p_c and the participation for the incorrect action p_i. In the iterations without spikes of the output neurons, normally the initial ones; the agent takes d_i as damage. Otherwise, the damage is calculated by

$$p_c(s_c, s_i) = \begin{cases} \dfrac{\max(0,\ \min(s_c, s_t)) - \max(0,\ \min(s_c, s_t)) + s_t}{2s_t} & s_c + s_i \le 2s_t \\ \dfrac{s_c}{s_c + s_i} & s_c + s_i > 2s_t \end{cases} \tag{12}$$

$$p_i(s_c, s_i) = 1 - p_c(s_c - p_i) \tag{13}$$

where s_t is the minimum "target" number of spikes. The purpose of s_t is to avoid assigning a too high or low fitness to agents that fire few spikes through their outputs. The agent takes damage at every time step and is given by

$$d(s_c, s_i) = d_c p_c(s_c, s_i) + d_i p_i(s_c, s_i) \tag{14}$$

Damaging is performed until the agent runs out of health points and 'dies'. Subsequently, the fitness score of the agent is calculated from the fitness function expressed in Eq. 11. The damage to the health points in a correct action d_c is 1, in an incorrect one d_i is 2. Therefore, correct actions result in a longer lifetime. The value for the minimum 'target' number of spikes s_t is 3 spikes.

In the control task of cart-pole balancing, the behavior of the mutable environment is different. A new environment is presented to the agent either after its failure or after the maximum number of environment iterations is reached. Moreover, the agents do not have health points. The fitness score is the normalization of the number of iterations that the agent survived after all environment conditions were executed.

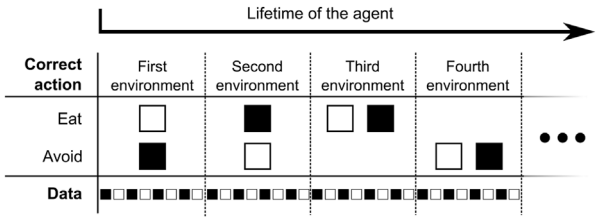


FIGURE 3
Example of the food foraging environment and how it progresses through the lifetime of the agent in a generation. The eight food samples per environment are illustrative. Our experiment uses four.

TABLE 1 Correct actions for all combinations of input food color and edible food in the food foraging task.

Food foraging environment conditions				
Edible	**Black**	**White**	**None**	**Both**
Input				
Black	Eat	Avoid	Avoid	Eat
White	Avoid	Eat	Avoid	Eat

4.3.1 Food foraging

The agent in the food foraging environment possesses just one light sensor for identifying the food "in front of it." There are two types of food: edible and poisonous. As such, food is represented in two colors: black and white. The environment changes by randomly defining which food color is edible or poisonous. In this environment, the agent can act in two ways: eating or avoiding the food. The sample has a predefined time of exposure to the agent. An action is performed after the first spike and it continues for every time step in the environment simulation. After this exposure time, the food is replaced by a new one. The agent can only discover whether it is exposed to an edible or poisonous food by interacting with it. An incorrect action is defined as eating poisonous food, or avoiding edible food, while a correct action is defined as eating edible food or avoiding poisonous one. If the agent makes an incorrect action, it receives a penalty signal, from which the agent should learn over the generations that it represents pain, revulsion, or hunger. If the agent makes a correct action, it receives a reward signal, from which it should learn that it represents the pleasure of eating delicious food or recognizing that the food is poisonous. In Figure 3, the food foraging environment is illustrated, how the environment changes and provides new food samples. In our experiment, the change of the environment occurs after presenting four food samples to the agent. The first food sample type is chosen randomly and alternates in every sample change. In Table 1, the four combinations of edible and poisonous food for the white and black ones are shown. To evolve the spiking neural network for the food foraging task, the parameters of the genetic algorithm are the following: the population size is set to 100 individuals, and the number of generations is set to 1,000. This task was chosen because of its simplicity. In particular, it allows a virtual wheeled robot to forage for food using proximity sensors, such as in the related work of Stanley et al. (2003).

4.3.2 Logic gates

In this environment, the mutable environmental state is a two-input logic gate. The environment provides the agent with two binary inputs, i.e., 0's and 1's. The agent's task is to predict the correct output for the current logic gate given the current input. Similar to the food foraging environment, it receives a reward signal if it is currently predicting the correct output, and a penalty signal if it is currently predicting the wrong output.

In order to measure the generalizing properties of agents, we use two different sets of environments: a training environment, which is used in calculating the fitness score while running the evolutionary algorithm, and a test environment which has a fully disjoint set of possible environmental states. A full overview of the logic gates found in both the training and the test environments, as well as the truth values for all input and output combinations, are found in Table 2 and Table 3. The evolution of the spiking neural network is performed by a population of 100 individuals through 1,000 generations.

4.3.3 Cart-pole balancing

The cart-pole balancing is a well-known control task used as a benchmark problem in reinforcement learning. In this environment, there is a cart that moves when a force is applied to the left or to the right every time step. In the middle of the cart, there is a vertical pole connected to a non-actuated joint. The goal of this environment is to maintain the pole balanced upright by controlling the forces that move the cart. Moreover, the cart cannot move beyond the limits of the track. The observations available to the controller are the cart position, the cart velocity, the pole angle, and the pole angular velocity.

For training, we use poles of different sizes, which are 0.5 (default), 0.3, and 0.7. For testing, the sizes are 0.4, and 0.6. Those pole sizes are depicted in the Supplementary Material. Each size can run up to 200 environment iterations and it is repeated three times during training for promoting stable controllers. If there are no more environment iterations or the pole falls, the cart-pole environment restarts with the next pole size while using the same SNN or finishes when all pole sizes were executed. The fitness score is calculated using the number of iterations the pole kept balanced. Subsequently, it is normalized to values between 0 and 1. The evolution for this task occurs with a population size of 256 during 500 generations.

TABLE 2 Truth table showing the correct output for each training logic gate.

Training logic gates

Input		A	B	NOT A	NOT B	Only 0	Only 1	XOR	XNOR
A	B								
0	0	0	0	1	1	0	1	0	1
0	1	0	1	1	0	0	1	1	0
1	0	1	0	0	1	0	1	1	0
1	1	1	1	0	0	0	1	0	1

TABLE 3 Truth table showing the correct output for each testing logic gate.

Test logic gates

Input		AND	NAND	OR	NOR
A	B				
0	0	0	1	0	1
0	1	0	1	1	0
1	0	0	1	1	0
1	1	1	0	1	0

4.4 Data representation

The data type in a spiking neural network is a binary time series or a spike train. Because the agent senses and acts in the environment, such data must be converted from the sensors and to the actuators. The flow of spikes over time can be quantified as firing rate, which corresponds to a frequency, or the number of spikes per second. The firing rate is the data representation that is converted as inputs and outputs for the SNN. However, the input firing rate must be within a minimum and a maximum value. In our experiments, we use the value range [5*Hz*, 50*Hz*]. The minimum and maximum value of the firing rate are simplified to a real number range [0, 1]. It is preferable that the data from the sensors has also a minimum and a maximum value. Otherwise, it will be necessary to clip sensory values or map the values to a desirable range.

In the binary classification tasks, all inputs and outputs are binary. Therefore, the minimum and maximum values for the input firing rate stand for, respectively, 0 and 1, or *False* and *True*. To avoid having a predefined threshold firing rate for the output neurons, we opt to have two output neurons for one binary value. The neuron with the highest firing rate within the actuator time window is the one defining the binary output value. If these two output neurons have the same firing rate, then the last one with

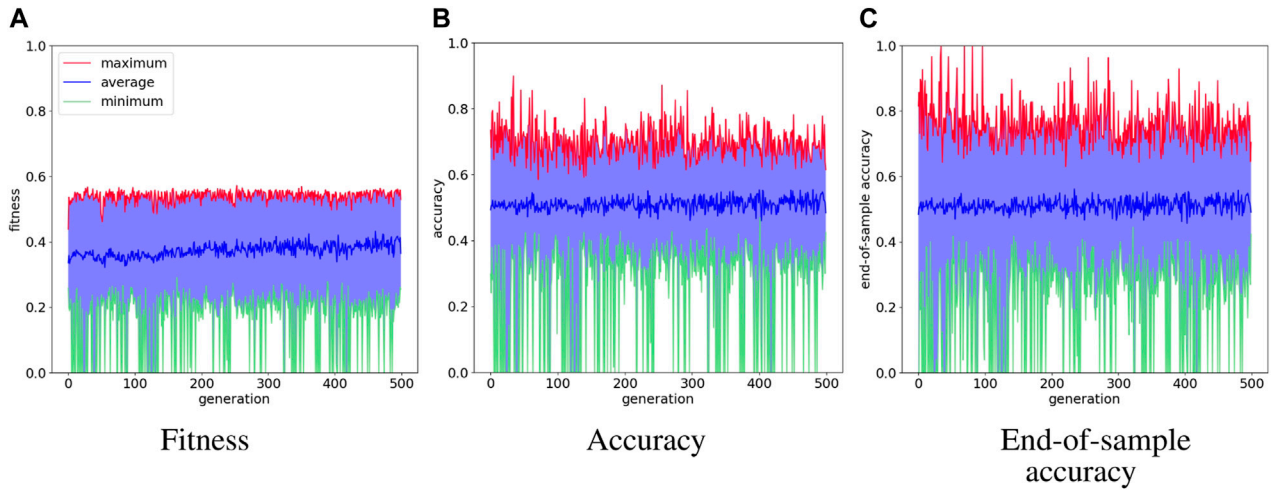


FIGURE 4
Evolution history of food foraging environment showing the average, minimum and maximum per generation.

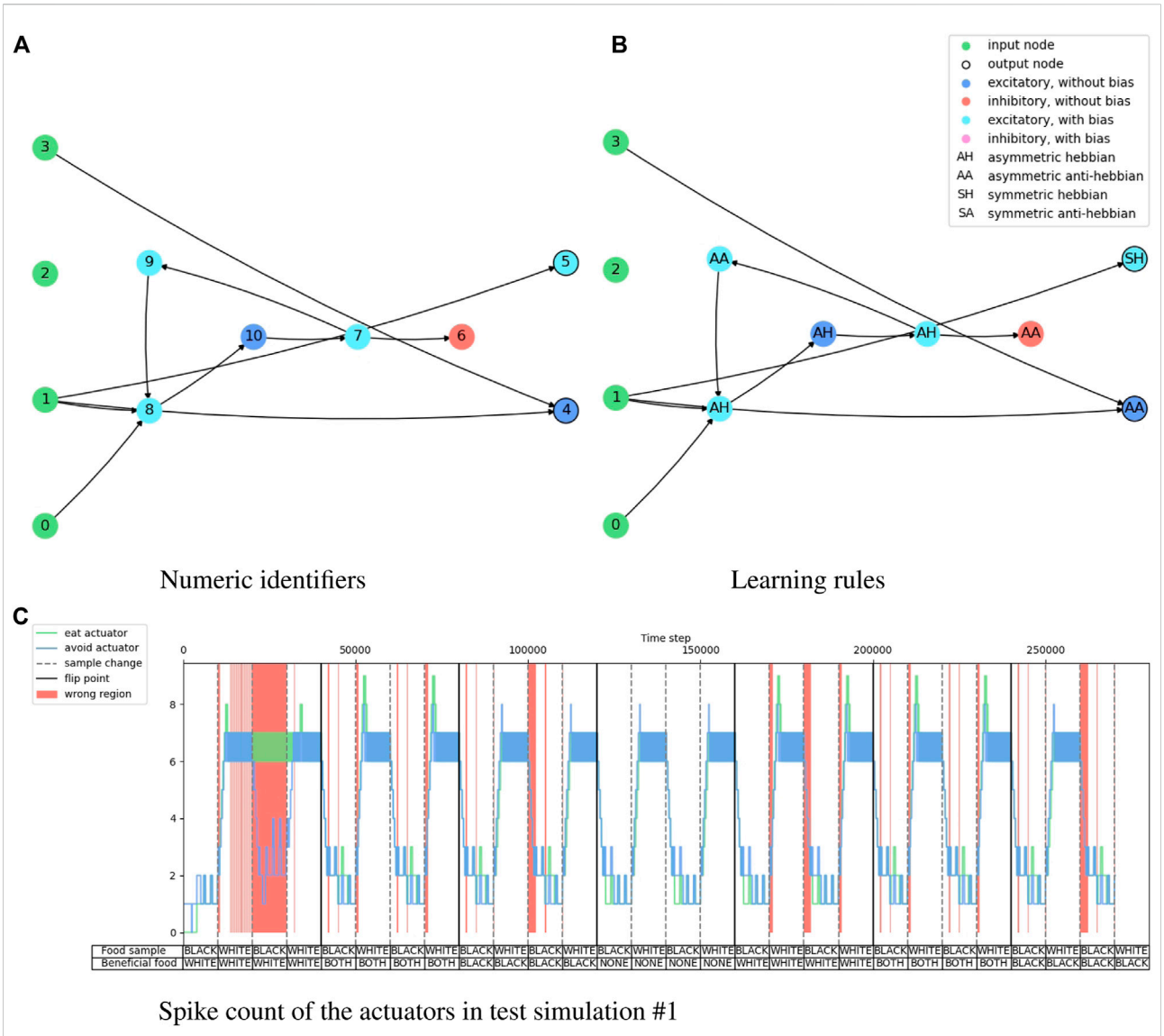


FIGURE 5
Illustration of the network topology and behavior of the highest accuracy agent in the food foraging task. The one-hot encoded input sample goes into nodes 0 and 1, the reward signal in node 2, and the penalty signal goes into node 3. Node 4 is the output for the "eat" actuator and node 5 is the output for the "avoid" actuator.

the highest value is selected. We also decided to have the same "two neurons-one binary value" strategy with the inputs, which consists of 0 or *False* being 01 in one-hot encoding, then (*low*, *high*) in firing rate, while 1 or *True* is 10 in one-hot encoding, so the firing rate is (*high*, *low*).

For the cart-pole control task, the inputs are real numbers, and the left and right actions are represented as two output neurons, similar to the outputs of the binary classification tasks. In this environment, the inputs are the cart position, cart velocity, pole angle, and pole angular velocity. Because we infer that real numbers converted to the firing rate of one neuron can be difficult to deal with in an adaptive spiking neural network (as also mentioned in Ref. (Pontes-Filho and Liwicki, 2019)), we decided to have three neurons for each input. The firing rate of the three neurons is similar to the sensitivity for the light spectrum of the three cone cells in the human eye (Bowmaker and Dartnall, 1980). We use the sigmoid function (Han and Moraga, 1995) for neurons #1 and #3 and a normalized version of the Gaussian function (Patel and Read, 1996) for neuron #2. The sigmoid equation is

$$\mathcal{F}_{sigmoid}(x \mid \omega, z, h, l) = \frac{h}{1 + e^{-\omega(x-z)}} + l, \tag{15}$$

where x is the observation value from the environment, ω is the weight that adjusts the smoothness of the interval between 0 and

TABLE 4 Test simulations of the highest accuracy agent in the food foraging experiment. "Acc." stands for accuracy and "EOS Acc." for end-of-sample accuracy.

Food foraging test simulations

#	Acc. (%)	EOS Acc. (%)	Input order	Environment order
1	88.0	92.6	black, white	white, both, black, none
2	90.6	100	white, black	white, none, both, black
3	91.3	100	black, white	white, both, none, black
4	85.4	92.3	white, black	white, black, both, none
5	89.5	96.3	white, black	both, none, white, black
6	89.2	100.0	black, white	both, white, black, none
7	87.7	92.6	black, white	white, black, none, both
8	84.9	92.6	black, white	black, both, white, none
9	89.8	100	black, white	white, black, both, none
10	88.4	92.6	white, black	black, none, white, both
Avg	88.4	95.9	n/a	

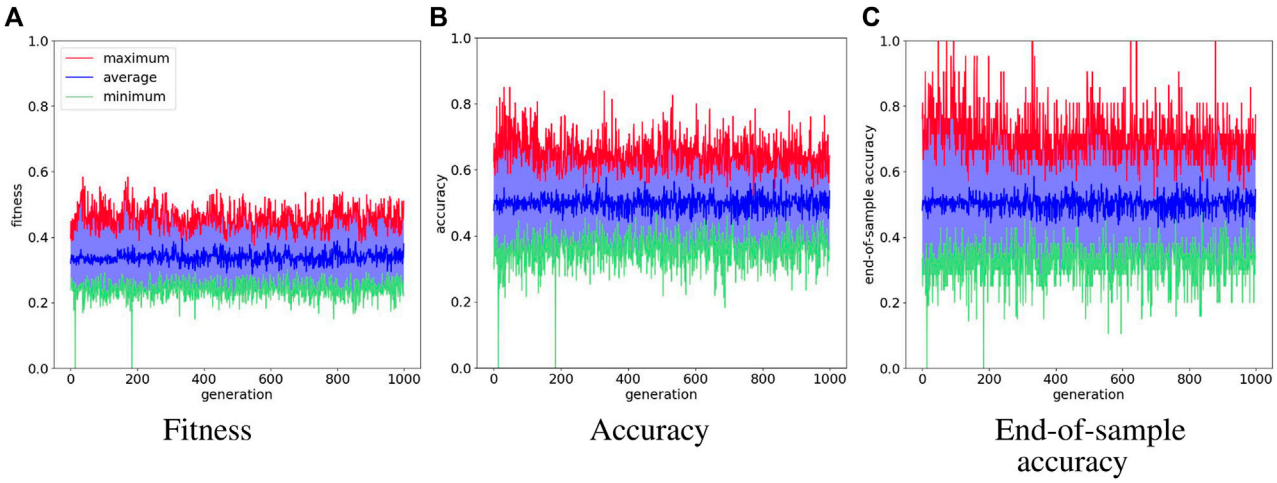


FIGURE 6
Evolution history of logic gate environment showing the average, minimum and maximum per generation.

1, z is the shift coefficient to adjust the function on the horizontal axis, h is the highest firing rate possible applied to an input neuron, and l is the lowest firing rate possible. The Gaussian function for converting observation value to firing rate is expressed by

$$\mathcal{F}_{Gaussian}(x \mid \mu, \sigma, h, l) = he^{\frac{-\left((x-\mu)^2\right)}{2\sigma^2}} + l, \tag{16}$$

where μ is the mean and σ is the standard deviation. We replace $\frac{1}{\sigma\sqrt{2\pi}}$ in the original Gaussian function to h because, in this way, we can define the highest firing rate when the observation value is the mean. Neurons #1 and #3 use $\mathcal{F}_{sigmoid}$, while neuron #2 uses $\mathcal{F}_{Gaussian}$. The parameters and the figures with the illustration of those equations are included in the Supplementary Material.

5 Results

The evolution of the spiking neural networks in NAGI is evaluated with fitness score, accuracy, and end-of-sample accuracy for the binary classification tasks, which are food foraging and logic gate. The accuracy is measured at every time step of the simulation. The end-of-sample accuracy stands for the accuracy measured in the last time step of a

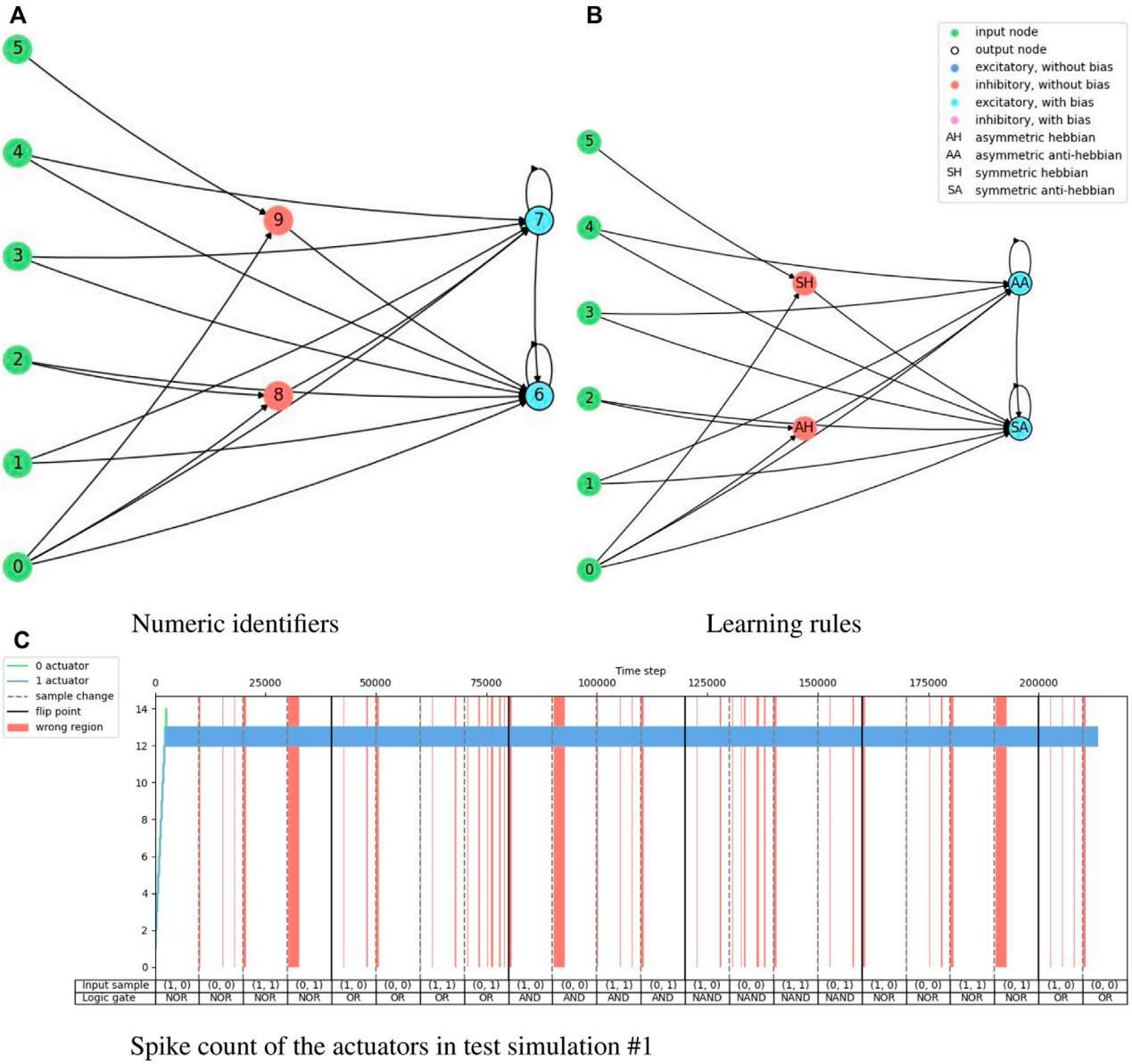


FIGURE 7
Illustration of the network topology and behavior of the highest training accuracy agent in the logic gate task. The one-hot encoded input sample "**(A)**" goes into nodes 0 and 1, the one-hot encoded input sample "**(B)**" goes into nodes 2 and 3, the reward signal goes into node 4, and the penalty signal into node 5. Node 6 is the output for the "0" actuator and node 7 is the output for the "1" actuator.

sample. The assessment performed for the control task with cart-pole balancing is done with the fitness score. We test the best performing agent in a task with ten simulations where their details are also provided.

Figure 4 shows the evolution history of the food foraging task. The average fitness score has a slight increase, but the maximum fitness score does not follow this trend. The accuracy and end-of-sample accuracy have high variation with their maximum values, but they consist of high accuracies. Moreover, some early generations register 100% end-of-sample accuracy. The three measurements do not improve through the generations. However, good solutions are already found in the first generation. Therefore, this is an easy task that requires a small SNN. For test simulations, we select the individual with the highest accuracy, which is found in generation number 34 and has an accuracy of 89.8%. Its fitness score is 0.541395 and its end-of-sample accuracy is 100%. Its topology is shown in Figure 5. Paying attention to this topology, the hidden nodes are not

TABLE 5 Test simulations of the highest training accuracy agent in the logic gate experiment. "Acc." stands for accuracy and "EOS Acc." for end-of-sample accuracy.

Logic gate test simulations

#	Acc. (%)	EOS Acc. (%)	Input order (A, B)	Environment order
1	89.8	100	(1, 0), (0, 0), (1, 1), (0, 1)	NOR, OR, AND, NAND
2	85.2	95.2	(1, 1), (1, 0), (0, 0), (0, 1)	OR, NOR, NAND, AND
3	86.0	100	(1, 0), (1, 1), (0, 1), (0, 0)	NOR, OR, AND, NAND
4	85.9	95.2	(0, 0), (1, 1), (0, 1), (1, 0)	NAND, AND, OR, NOR
5	79.9	85.7	(0, 0), (0, 1), (1, 0), (1, 1)	NAND, AND, NOR, OR
6	88.8	100	(1, 0), (0, 0), (1, 1), (0, 1)	AND, NAND, OR, NOR
7	85.1	90.5	(0, 0), (1, 1), (1, 0), (0, 1)	OR, NOR, NAND, AND
8	84.8	90.5	(1, 1), (0, 1), (0, 0), (1, 0)	NOR, NAND, OR, AND
9	83.7	85.7	(0, 0), (1, 0), (0, 1), (1, 1)	NAND, NOR, OR, AND
10	88.5	100	(1, 1), (1, 0), (0, 0), (0, 1)	NOR, AND, OR, NAND
Avg	85.7	94.2	n/a	

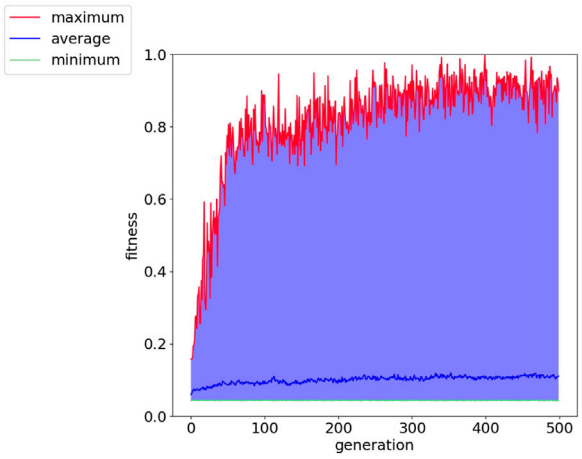


FIGURE 8
Fitness history of cart-pole balancing environment showing the average, minimum and maximum per generation.

TABLE 6 Test simulations of the highest fitness agent in the cart-pole balancing experiment.

Cart-pole balancing test simulations

#	Fitness	# Steps 0.4	# Steps 0.6	Environment order
1	1.000	200	200	0.4, 0.6
2	1.000	200	200	0.4, 0.6
3	1.000	200	200	0.6, 0.4
4	0.943	200	177	0.4, 0.6
5	0.800	154	166	0.6, 0.4
6	0.792	179	138	0.4, 0.6
7	0.835	200	134	0.6, 0.4
8	0.845	200	138	0.4, 0.6
9	0.873	200	149	0.6, 0.4
10	0.720	88	200	0.6, 0.4
Avg	0.874	178.3	171.5	n/a

needed. They form a loop that does not connect with the output nodes. The topology summarizes in one of the one-hot encoded input nodes (node 1) connecting to the two output nodes. Then, the node with the penalty signal (node 3) connects only with the node for the "eat" actuator (node 4). The behavior of the network is illustrated in Figure 5C. The topology of the network indicates that the two output neurons have the same data input from node 1, but the neuron for "avoid" action has a bias, which gives it a small excitatory current. If "avoid" is the wrong action, the penalty input signal from node 3 excites the output neuron for the "eat" action. This is how the spiking neural network decides the actions from "understanding" the feedback of the environment given by the penalty input signal. The result of the ten test simulations is presented in Table 4.

Figure 6 shows the training results of the logic gate task and it includes the test of the maximum individual of the measurement in every generation. The fitness score, accuracy, and end-of-sample accuracy maintain average values with high variation. However, the evolution of the agents in the logic gate task is similar to the one in the food foraging. The early generations already contain good spiking neural networks for the task. The best-performing agent is selected from the accuracy measurement. This individual is in generation 48 and has an accuracy of 85.0%. Its fitness score is 0.4421625 and its end-of-sample accuracy is 100%. The topology of this spiking neural

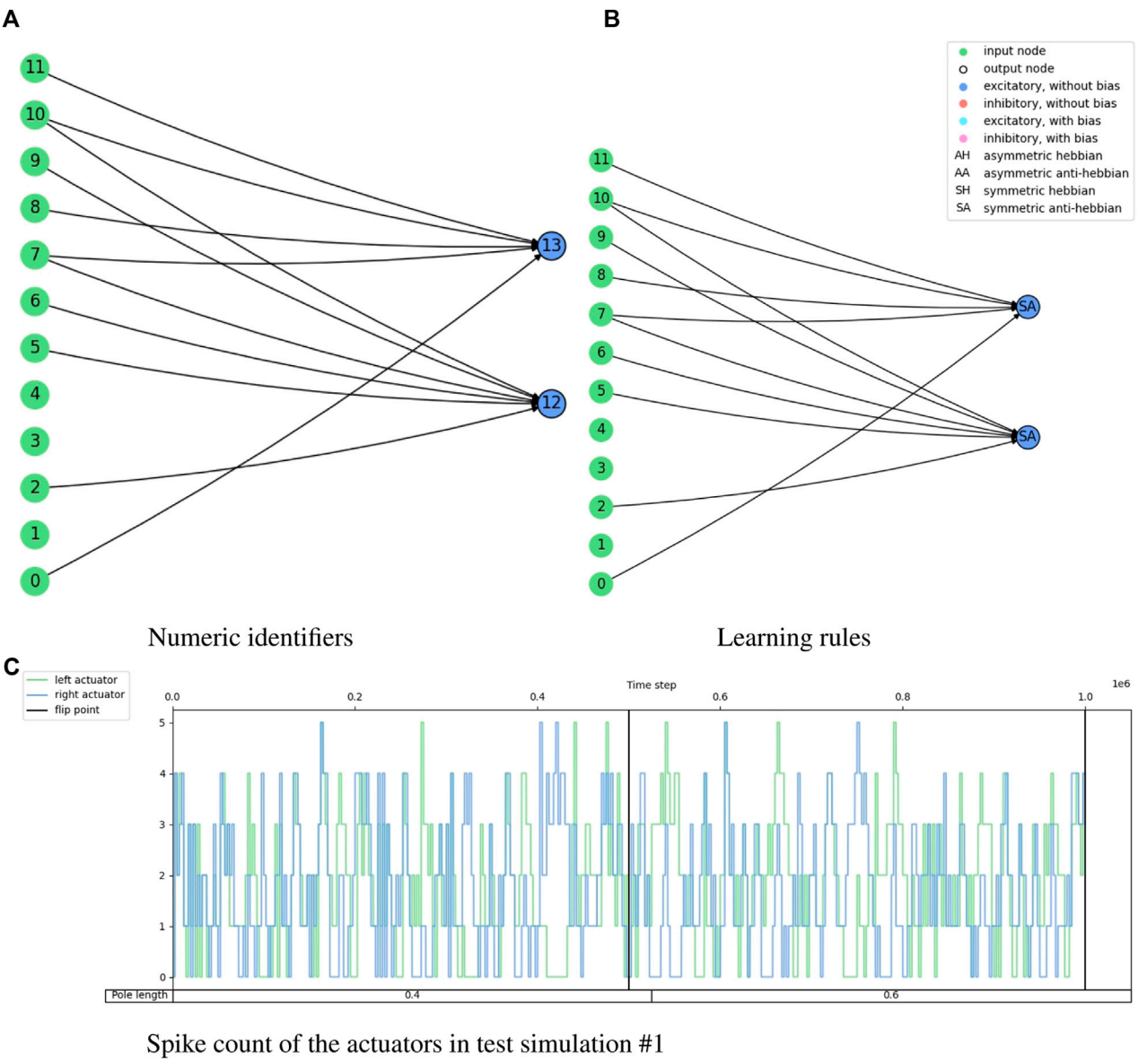


FIGURE 9
Illustration of the network topology and behavior of the highest training fitness agent in the cart-pole balancing experiment. The 3-tuple of input nodes (0, 1, 2) gets the converted firing rate from the observation of the cart position, (3, 4, 5) from the cart velocity, (6, 7, 8) from the pole angle, and (9, 10, 11) from the pole angular velocity. Node 12 is the output for the 'left' action and node 13 is the output for the "right" action.

network is shown in Figure 7. Its behavior is shown in Figure 7C. Even though we have trained with a "confidence" factor in the fitness function, the spike counts are still with almost the same values. Table 5 contains the accuracy and end-of-sample accuracy of ten test simulations, which indicates that the SNN can be general to reproduce the behavior of logic gates without being trained to them.

Figure 8 shows the fitness score history through the evolution for the cart-pole balancing task. This task is the one with the highest difficulty to find a good genome for the adaptive spiking neural network. It can be noted that the fitness score improves through the generations. The maximum fitness score in a generation goes from around 0.16 in the first generation to 0.99944 in generation number 399. Such an individual is the one selected for the test simulations. Its topology is illustrated in Figure 9 and the spike counts of the actuators for "left" and "right" actions are shown in Figure 9C. The spiking neural network has no

hidden neurons. Therefore, the SNN works as an input selection for the output neurons. The result of the ten test simulations is presented in Table 6. When the pole is balanced for more than 100 iterations, the controller is considered successful.

6 Discussion and conclusion

We successfully solved all three presented tasks with the NAGI framework. The spiking neural networks found showed generality to the binary classification tasks, even to unseen conditions in the case of the emulation of logic gates. The neuroevolution produced rather simple topologies for the SNNs. We infer that binary classification is easy due to the binary performance feedback. For further research, multi-class classification is considered.

The cart-pole balancing task was successfully solved without any hidden neurons. The conversion of one observation into three input neurons is used to avoid the requirement of weight fine-tuning due to small differences in firing rate and also to the assumption that Hebbian plasticity works better with binary data (active and inactive) (Pontes-Filho and Liwicki, 2019). With such a conversion, the SNN became an input selection.

The topologies for the three tasks caught our attention because almost all output excitatory neurons were anti-Hebbian, and the two inhibitory hidden neurons in the logic gate solution have Hebbian neuroplasticity. Our initial hypotheses were that excitatory neurons mainly have Hebbian learning rules, and inhibitory neurons are anti-Hebbian. That was the reason for having different probabilities for anti-Hebbian and Hebbian learning rules depending on the type of the neurotransmitter when adding a new neuron through mutation.

Even though there is elitism, the performance measurements are unstable through generations. This is a demonstration of the randomness in the initialization of the weights, and input and environment order. This can be perceived in the results of the ten test simulations of the three tasks.

For future work, we plan to attempt more challenging tasks. If there is a failure in executing the task, the constraints imposed on NAGI can be eased. A major constraint is that one neuron has one plasticity rule for all dendrites. Maybe its removal can simplify issues in difficult tasks. This constraint was intended to reduce the dimensionality of the search space in the neuroevolution and an assumption that the dendrites in the same neuron adapt under one learning rule. This modification is also aligned with the work of Najarro and Risi, (2020), which has meta-learning properties for more difficult control tasks than the cart-pole balancing, such as top-down car racing and quadruped walk. Another opportunity is the addition of curriculum learning (Bengio et al., 2009; Narvekar et al., 2020) for increasing the complexity of the task while the agent becomes better over the generations.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

SP-F had the main idea, supervised most part of the work, implemented additional experiments, and wrote the initial draft of the manuscript. KO contributed to the writing of the initial draft and performed most of the experimental work as part of his master's thesis while being mainly supervised by SP-F. AY, MR, PH, and SN co-supervised the work. All authors reviewed the manuscript.

Funding

This work was partially funded by the Norwegian Research Council (NFR) through their IKTPLUSS research and innovation action under the project Socrates (grant agreement 270961).

Acknowledgments

This manuscript includes materials from the Master's thesis of Olsen, (2020). Its preprint is available in Ref. (Pontes-Filho et al., 2022).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2022.1007547/full#supplementary-material

References

Ardiel, E. L., and Rankin, C. H. (2010). An elegant mind: Learning and memory in caenorhabditis elegans. *Learn. Mem.* 17, 191–201. doi:10.1101/lm.960510

Bengio, Y., Louradour, J., Collobert, R., and Weston, J. (2009). "Curriculum learning," in Proceedings of the 26th annual international conference on machine learning, 41–48.

Betts, J. G., Young, K. A., Wise, J. A., Johnson, E., Poe, B., Kruse, D. H., et al. (2013). *Anatomy and physiology*. Houston, TX: OpenStax. Available at: https://openstax.org/details/books/anatomy-and-physiology.

Bowmaker, J. K., and Dartnall, H. J. (1980). Visual pigments of rods and cones in a human retina. *J. Physiology* 298, 501–511. doi:10.1113/jphysiol.1980.sp013097

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., et al. (2016). Openai gym. *arXiv preprint arXiv:1606.01540*

Camuñas-Mesa, L. A., Linares-Barranco, B., and Serrano-Gotarredona, T. (2019). Neuromorphic spiking neural networks and their memristor-cmos hardware implementations. *Materials* 12, 2745. doi:10.3390/ma12172745

Crosby, M., Beyret, B., and Halina, M. (2019). The animal-ai olympics. *Nat. Mach. Intell.* 1, 257. doi:10.1038/s42256-019-0050-3

Diehl, P., and Cook, M. (2015). Unsupervised learning of digit recognition using spike-timing-dependent plasticity. *Front. Comput. Neurosci.* 9, 99. doi:10.3389/fncom.2015.00099

Doncieux, S., Bredeche, N., Mouret, J.-B., and Eiben, A. E. G. (2015). Evolutionary robotics: What, why, and where to. *Front. Robot. AI* 2, 4. doi:10.3389/frobt.2015.00004

Funes, P., and Pollack, J. (1998). Evolutionary body building: Adaptive physical designs for robots. *Artif. Life* 4, 337–357. doi:10.1162/106454698568639

Gaier, A., and Ha, D. (2019). "Weight agnostic neural networks," in *Advances in Neural Information Processing Systems*. Editors H. Wallach, H. Larochelle, A. Beygelzimer, F. D'. Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc.) 32. Available at: https://proceedings.neurips.cc/paper/2019/file/e98741479a7b998f88b8f8c9f0b6b6f1-Paper.pdf.

Han, J., and Moraga, C. (1995). "The influence of the sigmoid function parameters on the speed of backpropagation learning," in Proceedings of the International Workshop on Artificial Neural Networks: From Natural to Artificial Neural Computation (London, UK, UK: Springer-Verlag), 195–201.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Holland, J. H. (1992). Genetic algorithms. *Sci. Am.* 267, 66–72. doi:10.1038/scientificamerican0792-66

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 14, 1569–1572. doi:10.1109/tnn.2003.820440

Kulik, Y., Jones, R., Moughamian, A. J., Whippen, J., and Davis, G. W. (2019). Dual separable feedback systems govern firing rate homeostasis. *Elife* 8, e45717. doi:10.7554/elife.45717

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J. (2017). Building machines that learn and think like people. *Behav. Brain Sci.* 40, e253. doi:10.1017/s0140525x16001837

Langton, C. (2019). *Artificial life: Proceedings of an interdisciplinary workshop on the synthesis and simulation of living systems*. London: Routledge.

Li, Y., Zhong, Y., Zhang, J., Xu, L., Wang, Q., Sun, H., et al. (2014). Activity-dependent synaptic plasticity of a chalcogenide electronic synapse for neuromorphic systems. *Sci. Rep.* 4, 4906. doi:10.1038/srep04906

Liu, Y.-H., and Wang, X.-J. (2001). Spike-frequency adaptation of a generalized leaky integrate-and-fire model neuron. *J. Comput. Neurosci.* 10, 25–45. doi:10.1023/a:1008916026143

Mautner, C., and Belew, R. K. (2000). Evolving robot morphology and control. *Artif. Life Robot.* 4, 130–136. doi:10.1007/bf02481333

Nadji-Tehrani, M., and Eslami, A. (2020). A brain-inspired framework for evolutionary artificial general intelligence. *IEEE Trans. Neural Netw. Learn. Syst.* 31, 5257–5271. doi:10.1109/tnnls.2020.2965567

Najarro, E., and Risi, S. (2020). "Meta-Learning through Hebbian Plasticity in Random Networks," in *Advances in Neural Information Processing Systems*. Editors H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Curran Associates, Inc.) 33, 20719–20731. Available at: https://proceedings.neurips.cc/paper/2020/file/ee23e7ad9b473ad072d57aaa9b2a5222-Paper.pdf.

Narvekar, S., Peng, B., Leonetti, M., Sinapov, J., Taylor, M. E., and Stone, P. (2020). Curriculum learning for reinforcement learning domains: A framework and survey. *J. Mach. Learn. Res.* 21 (181), 1–50.

Olsen, K. (2020). *Neuroevolution of artificial general intelligence*. Oslo: University of Oslo.

Patel, J. K., and Read, C. B. (1996). *Handbook of the normal distribution*, 150. CRC Press.

Pontes-Filho, S., and Liwicki, M. (2019). "Bidirectional learning for robust neural networks," in 2019 International Joint Conference on Neural Networks (IEEE), IJCNN '19, 1–8.

Pontes-Filho, S., and Nichele, S. (2019). A conceptual bio-inspired framework for the evolution of artificial general intelligence. *arXiv preprint arXiv:1903.10410*

Pontes-Filho, S., Olsen, K., Yazidi, A., Riegler, M., Halvorsen, P., and Nichele, S. (2022). Towards the neuroevolution of low-level artificial general intelligence. *arXiv preprint arXiv:2207.13583*

Randi, F., and Leifer, A. M. (2020). Measuring and modeling whole-brain neural dynamics in caenorhabditis elegans. *Curr. Opin. Neurobiol.* 65, 167–175. doi:10.1016/j.conb.2020.11.001

Reed, S., Zolna, K., Parisotto, E., Colmenarejo, S. G., Novikov, A., Barth-Maron, G., et al. (2022). A generalist agent. *arXiv preprint arXiv:2205.06175*

Risi, S., and Stanley, K. O. (2010). "Indirectly encoding neural plasticity as a pattern of local rules," in International Conference on Simulation of Adaptive Behavior. Springer, 533–543.

Risi, S. (2021). The future of artificial intelligence is self-organizing and self-assembling. Available at: sebastianrisi.com

Shen, Z., Liu, J., He, Y., Zhang, X., Xu, R., Yu, H., et al. (2021). Towards out-of-distribution generalization: A survey. *arXiv preprint arXiv:2108.13624*

Stanley, K. O., Bryant, B. D., and Miikkulainen, R. (2003). Evolving adaptive neural networks with and without adaptive synapses. In The 2003 Congress on Evolutionary Computation (IEEE), 2557–2564.

Stanley, K. O., Clune, J., Lehman, J., and Miikkulainen, R. (2019). Designing neural networks through neuroevolution. *Nat. Mach. Intell.* 1, 24–35. doi:10.1038/s42256-018-0006-z

Stanley, K. O., and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *Evol. Comput.* 10, 99–127. doi:10.1162/106365602320169811

Sukenik, N., Vinogradov, O., Weinreb, E., Segal, M., Levina, A., and Moses, E. (2021). Neuronal circuits overcome imbalance in excitation and inhibition by adjusting connection numbers. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2018459118. doi:10.1073/pnas.2018459118

Tavanaei, A., Ghodrati, M., Kheradpisheh, S. R., Masquelier, T., and Maida, A. (2019). Deep learning in spiking neural networks. *Neural Netw.* 111, 47–63. doi:10.1016/j.neunet.2018.12.002

Taylor, T. (2019). Evolutionary innovations and where to find them: Routes to open-ended evolution in natural and artificial systems. *Artif. Life* 25, 207–224. doi:10.1162/artl_a_00290

Thrun, S., and Pratt, L. (1998). "Learning to learn: Introduction and overview," in *Learning to learn* (Springer), 3–17.

Trappenberg, T. (2009). *Fundamentals of computational neuroscience*. Oxford: OUP Oxford.

Watson, R. A., Ficiei, S. G., and Pollack, J. B. (1999). "Embodied evolution: Embodying an evolutionary algorithm in a population of robots," in Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406).

Welleck, S. (2019). Evolving networks. Available at: wellecks.wordpress.com

Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nat. Commun.* 10, 3770–3777. doi:10.1038/s41467-019-11786-6

Zohora, F. T., Karia, V., Daram, A. R., Zyarah, A. M., and Kudithipudi, D. (2021). "Metaplasticnet: Architecture with probabilistic metaplastic synapses for continual learning," in 2021 IEEE International Symposium on Circuits and Systems (ISCAS) (IEEE).

frontiers | Frontiers in Robotics and AI

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/frobt.2022.1052998

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Dimitrije Marković,
Technical University Dresden, Germany
Tom Ziemke,
Linköping University, Sweden

*CORRESPONDENCE
Oscar Guerrero Rosado,
oscar.guerrerorosado@donders.ru.nl
Paul F. M. J. Verschure,
paul.verschure@donders.ru.nl

SPECIALTY SECTION
This article was submitted to Computational Intelligence in Robotics, a section of the journal Frontiers in Robotics and AI

RECEIVED 24 September 2022
ACCEPTED 17 November 2022
PUBLISHED 02 December 2022

CITATION
Rosado OG, Amil AF, Freire IT and Verschure PFMJ (2022), Drive competition underlies effective allostatic orchestration.
Front. Robot. AI 9:1052998.
doi: 10.3389/frobt.2022.1052998

Drive competition underlies effective allostatic orchestration

Oscar Guerrero Rosado*, Adrian F. Amil, Ismael T. Freire and Paul F. M. J. Verschure*

Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, Netherlands

Living systems ensure their fitness by self-regulating. The optimal matching of their behavior to the opportunities and demands of the ever-changing natural environment is crucial for satisfying physiological and cognitive needs. Although homeostasis has explained how organisms maintain their internal states within a desirable range, the problem of orchestrating different homeostatic systems has not been fully explained yet. In the present paper, we argue that attractor dynamics emerge from the competitive relation of internal drives, resulting in the effective regulation of adaptive behaviors. To test this hypothesis, we develop a biologically-grounded attractor model of allostatic orchestration that is embedded into a synthetic agent. Results show that the resultant neural mass model allows the agent to reproduce the navigational patterns of a rodent in an open field. Moreover, when exploring the robustness of our model in a dynamically changing environment, the synthetic agent pursues the stability of the self, being its internal states dependent on environmental opportunities to satisfy its needs. Finally, we elaborate on the benefits of resetting the model's dynamics after drive-completion behaviors. Altogether, our studies suggest that the neural mass allostatic model adequately reproduces self-regulatory dynamics while overcoming the limitations of previous models.

KEYWORDS

self-regulation, homeostasis, allostatic control, attractor model, need-based behavior, control theory

Introduction

As 19th-century physiologists Claude Bernard and Ivan Pavlov proposed, living systems are generally characterized by their ability to self-regulate (Bernard, 1865; Pavlov, 1955). Through self-regulation, an organism ensures its fitness by adjusting its inner processes relative to external perturbations (Papies and Aarts, 2016), in turn, assisting self-maintenance (i.e., Autopoiesis) (Maturana and Varela, 1991). One aspect of self-regulation is homeostasis, which describes the process of maintaining an internal state within a desirable range as proposed by Cannon (Cannon, 1939). Once external perturbations produce a deviation from the desirable range, a homeostatic error arises, driving a proportional error-correcting response to restore balance in the system. This process would be later formalized as a feedback control loop by Norbert Wiener (Wiener, 1948), father of cybernetics. However, the homeostatic control of a

single need is insufficient to ensure fitness since living organisms need to maintain a rather large set of internal needs with dynamically varying priorities from electrolysis to temperature and oxygenation. To overcome this and other gaps of homeostasis in explaining self-regulation, the notion of allostasis aims to capture the stability of the integrated self rather than its parts.

Allostasis transcends the constancy imposed by homeostasis by aligning all organism's internal parameters with the environmental demands and opportunities (Sterling, 1988). In its broad conceptualization, allostasis targets stability by integrating many dynamic regulatory principles (Sterling, 2020). For this reason, we differentiate three complementary and coupled levels of allostasis. First, allostatic orchestration allows the agent to rapidly rank its internal needs based on priorities, urgencies, and opportunities. Second, predictive allostasis leverages environmental regularities to learn associations between external events and internal states, supporting anticipatory allostatic orchestration and future homeostatic risks. Finally, contextual allostasis benefits from goal-oriented learning mechanisms to enrich self-regulatory strategies with spatio-temporal information, in turn facilitating anticipatory behavioral strategies. Hence, we consider that in the mammalian brain allostatic regulation of action is organized in a multi-layered architecture following the Distributed Adaptive Control (DAC) theory (Verschure, 2012, 2016).

While nowadays, allostasis is increasing in popularity (2010: 2,584 citations *versus* 2021:12549 citations, resource: Scopus), most of the latest computational models on allostasis focus on exploring the advantages of its predictive component (Sterling, 2012). Paradoxically, some models return to the one-single-need problem of homeostasis (Tschantz et al., 2022). In addition, the concept of allostasis has not explained the computational mechanisms by which individuals achieve stability by orchestrating different homeostatic systems. The divergent modeling approaches adopted by the few computational studies addressing allostasis demonstrate the lack of consensus when determining the fundamental principles behind allostatic orchestration.

In 2010, Sanchez-Fibla et al. developed what, to our knowledge, is the first computational model of allostatic orchestration (Sanchez-Fibla et al., 2010). The model emulated rodent behavior and physiological states in an open field test (Gould et al., 2009) with simulated and physical robots. This model proposed that the animal's behavior resulted from the interaction between two internal needs: Security, which is fulfilled in one of the arena's corners representing the rodent's home base, and arousal, which would be higher in the center of the arena given the maximum exposure of the animal at that location. However, although reproducing the animal's overall trajectory pattern and occupancy preferences, the model did not elaborate on the neuroscience supporting allostasis.

A more recent model bases the optimal selection of regulatory behaviors on maximizing a subsequent reward (Laurençon et al., 2021). This deployment of a reward-based allostatic model builds on the premises of homeostatic reinforcement learning (HRL) (Keramati and Gutkin, 2014). HRL represents a major refinement of traditional reinforcement learning theories grounding learning protocols on the individual's internal state. HRL successfully explains effects in animal behavior such as alliesthesia, namely, the fluctuations in reward value during resource acquisition (Cabanac, 1971). Still, although HRL leverages temporal and spatial information to improve the self-regulatory strategy, this approach makes a complementary learning process critical for solving the allostatic orchestration problem.

Finally, a third approach suggests that allostatic orchestration emerges from motivational conflict solved *via* attractor dynamics (Jimenez-Rodriguez et al., 2020).In this approach, the attractor dynamics implement competition through cross-inhibition (Usher and McClelland, 2001; Marshall et al., 2015). This framework supports the idea that attractor dynamics can underlie the optimal selection of regulatory behavior and explained their duration and latency. However, an explanation of how neural correlates of internal needs' implement such competing dynamics is unclear.

Grounding the design principles of a model of allostasis in state-of-the-art neuroscience constitutes a pending task for previous modeling approaches that have largely relied on algorithmic solutions. We suggest that novel approaches can overcome this challenge by focusing on the core behavior systems of the mammalian brain (Merker, 2013; Verschure, 2016). Interoception of physiological needs such as hydration, nutrition, thermoregulation, or sleep is generally attributed to distinct specialized hypothalamic nuclei (Strecker et al., 2002; Blouet and Schwartz, 2010; Nakamura, 2011; Zimmerman et al., 2017). In contrast, other psychological needs (e.g., social interaction) depend on more distributed brain networks (Lee et al., 2021). Importantly, recent studies suggest that these nuclei, and so the internal needs they represent, are not independent of each other but hold a competing relationship through inhibitory interactions (Burnett et al., 2016; Osterhout et al., 2022; Qian et al., 2022). This competition between internal drives could serve as the basis of allostatic orchestration by imposing a winner-take-all mechanism represented as attractor dynamics. Thus, irrelevant drives are suppressed, and the singleness principle of action is supported (Sherrington, 1906). This research literature suggests that an attractor-based approach is suitable for modeling allostasis.

In contexts where animals constantly self-regulate multiple internal needs, decision-making could be hampered by attractor forces sustained after drive-completion behaviors. Indeed, *in vivo* studies suggest that cortical areas involved in decision-making operate in a critical regime (close to a phase transition) (Ma et al.,

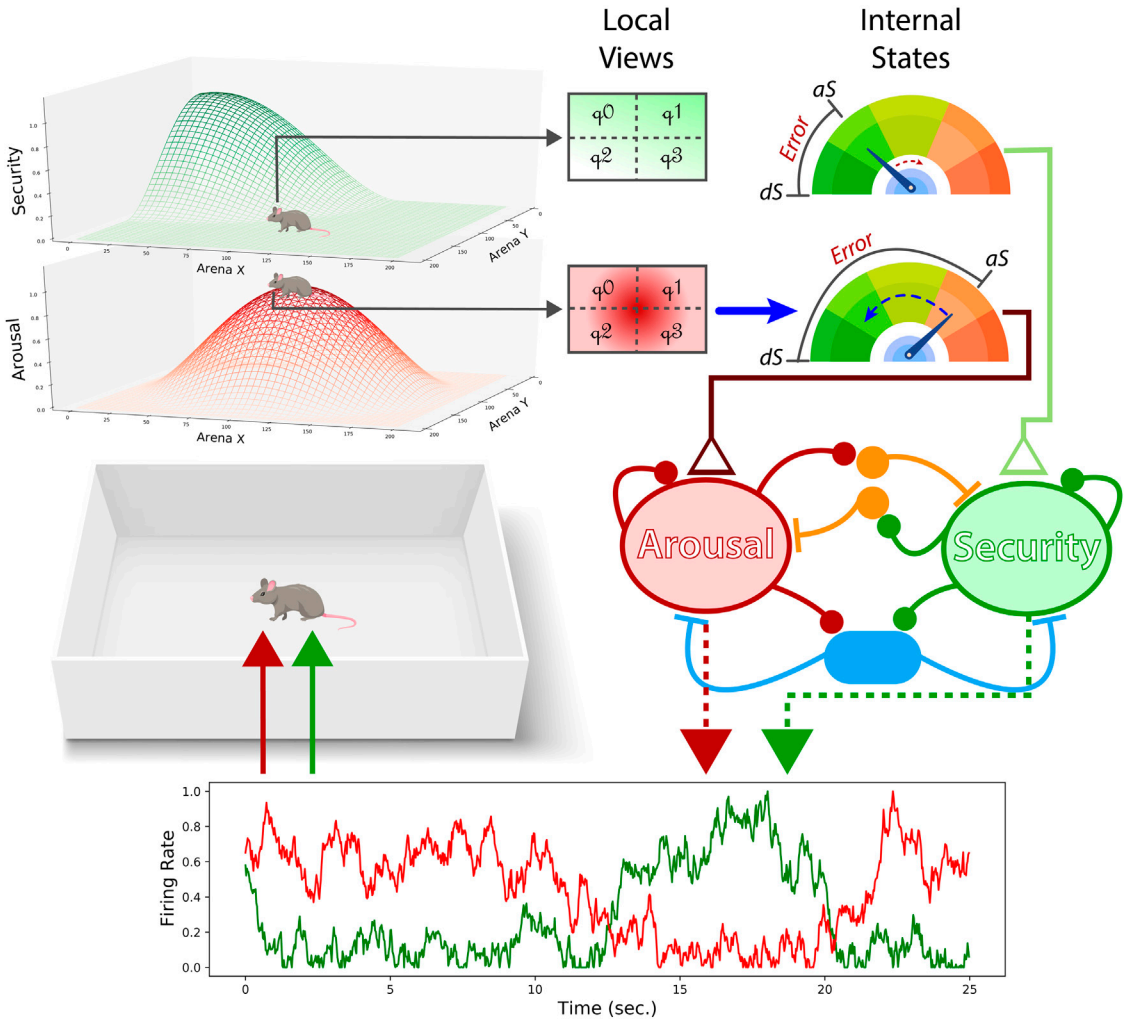


FIGURE 1
Performance of the neural mass allostatic model in an open field. Two gradients represent the areas where the two internal needs, arousal (red) and security (green), can be fulfilled (top-left). The agent partially observes those gradients through local sensation (top-middle). Local sensation allows the agent to adjust its actions to ascend/descend the gradients and detect when the observer's current position is in the vicinity of the resource area (i.e., the peak of the gradient). If the agent is not close enough to the resource area, the internal state related to that resource keeps declining, as the security actual state (aS) is doing in this figure (top-right). In contrast, if the agent occupies the peak of the gradient, the internal state approximates the desired state (dS), as shown with arousal. The aSs and dSs are compared, creating a homeostatic error (top-right) that will input their respective excitatory pools in the neural mass model (middle-right). The level of competition is defined by the ratio of mutual inhibition (orange interneurons) and shared feedback inhibition (blue interneurons). Finally, the firing rate of each excitatory population provides the agent with the corresponding drives given its internal states.

2019) that occasionally evolves to supercriticality (saturated population response), a regime that supports effective information transmission (Li et al., 2019). Nonetheless, how the system returns to criticality from a supercritical regime is poorly understood. The paraventricular hypothalamic nucleus (PVH) is a good candidate to inhibit the main interoceptive nuclei, recover their basal population activity after drive-completion behaviors, and create the initial conditions for the next cycle of allostatic orchestration. From one side, PVH mediates many diverse motivational functions, including thirst (Zimmerman et al., 2017), hunger (Blouet and Schwartz, 2010), and thermoregulation (Nakamura, 2011). Conversely, corticotrophin-releasing hormone (CRH) neurons in the PVH are suggested to be sensitive toward reward acquisition. Specifically, PVH CRH neurons get inhibited during drive-completion behaviors, representing a potential source of global inhibition to the rest of hypothalamic interoceptive nuclei (Yuan et al., 2019). In this research work, we model this form of decision reset by applying general inhibitory inputs to the excitatory populations once an internal need has been satisfied.

In the following sections, we will describe a novel neural mass model of allostatic orchestration grounded on the interoceptive mechanisms of the mammalian brain. As in (Sanchez-Fibla et al., 2010), the model will be embedded in an agent endowed with a so-called core behavior system (CBS) (Merker, 2013; Verschure, 2016). In other words, besides sensing its internal environment, a competence attributed to the hypothalamus (Strecker et al., 2002; Blouet and Schwartz, 2010; Nakamura, 2011; Zimmerman et al., 2017), the agent can orientate in an external environment and perform basic navigation based on an appetitive-aversive axis, cognitive functions attributed to the superior colliculus and the zona incerta-periaqueductal gray axis, respectively. The resultant model is tested in both static and dynamic environments. In the static condition, we aim to elucidate if our model can faithfully replicate both previous models of allostatic control and rodent behavior in an open field test (Gould et al., 2009). In the dynamic environment, we will further explore the robustness of our model when environmental opportunities to satisfy internal needs decrease over time. Finally, we will analyze how inducing subcritical dynamics after drive fulfillment facilitates switches in self-regulatory strategies.

Materials and methods

To better understand the potential of competing dynamics between internal drives in facilitating need orchestration and stability of the self, we built a novel allostatic model grounding its design on contemporary research literature. Consequently, a neural mass model was built incorporating two distinct populations sensitive to homeostatic markers while holding a competing relationship (Figure 1). Aiming for convergent validation, we equipped a synthetic agent with this biologically-constraint model and analyzed its ability to defend its internal states by navigating an external environment.

Homeostatic systems

Assuming that living organisms have interoceptive capabilities to assess their internal state, we conceptualized each internal need as a homeostatic mechanism. Here, actual and desired states are compared providing a measure of homeostatic error (that is, $hEi = |dSi - aSi|$). Importantly, homeostatic systems have no power to provide regulatory responses by themselves. Instead, they bias those behaviors by constantly feeding the neural mass allostatic model with homeostatic errors (Figure 1).

The transition between actual internal states responded to the following dynamical law,

$$aS_i(t) = aS_i(t-1) - dR_i(t) + rI_i(t) \tag{1}$$

where $aS_i(t-1)$ is the actual state of the homeostatic system i in the immediate previous timestep, $dR_i(t)$ corresponds to the decay rate applied to each homeostatic system at time t, and $rI_i(t)$ represents the resource impact at time t. Importantly, resource impact will be 0 unless the agent is in the vicinity of the resource. By applying different decay factors and resource impacts, internal states can evolve with different temporal dynamics.

This dynamical law, similarly applied in (Jimenez-Rodriguez et al., 2020), notably differs from the methods applied in (Sanchez-Fibla et al., 2010), where the agent drew its internal states directly from gradients mapped onto the two-dimensional arena.

The neural mass allostatic model

Wilson-Cowan equations modified as in (Amil and Verschure, 2021) were used to model two need-sensitive neural populations (Figure 1). This modification allowed to account for mutual and shared feedback inhibition held between the excitatory populations, as follows:

$$\begin{aligned}\tau\frac{dD_1(t)}{dt} = &-D_1 \\ &+ f\left(w_+D_1 + hE_1 - Qw_-D_2 - (1-Q)w_=f(D_1+D_2)\right) \\ &+ \sigma\xi(t)\end{aligned} \tag{2}$$

$$\begin{aligned}\tau\frac{dD_2(t)}{dt} = &-D_2 \\ &+ f\left(w_+D_2 + hE_2 - Qw_-D_1 - (1-Q)w_=f(D_1+D_2)\right) \\ &+ \sigma\xi(t)\end{aligned} \tag{3}$$

where $f(x)$ is the logistic $f - hE_i$ function,

$$f(x) = \frac{F_{max}}{1 + e^{\frac{-(x-\theta)}{k}}} \tag{4}$$

In these equations, τ is the time constant determining the timescale of population dynamics. D_i is the drive magnitude represented as the mean firing rate of the excitatory population i. w_+, w_-, and $w_=$ are the weights for recurrent connections within the excitatory population, mutual inhibition, and feedback inhibition, respectively. Q represents the mutual/feedback inhibition ratio, a variable that allows for inducing a controlled level of competition. σ and ξ are the variance and magnitude of Gaussian noise provided to the excitatory populations (See Supplementary Figures S1–S4 for Q and noise parameter search). Finally, $F_{\max}$, k, and θ are the maximum firing rate, gain, and threshold parameters of the $f - hE_i$ logistic curve respectively (Eq. 4).

Orientation

Methods enabling the orientation and navigation of the synthetic agent in the arena are based on (Sanchez-Fibla et al., 2010). In both conditions (static and dynamic environments), we served from three-dimensional gradients to represent the location of resources fulfilling specific internal needs. Notably, gradients are used solely to support navigation, and internal states are not directly linked to the agent's location in the arena, as in the original study. Instead, internal states follow their dynamics, as explained above. The decision to use gradient-based methods for navigation is supported by representations of future navigational goals in the orbitofrontal cortex (Basu et al., 2021).

We implemented a partially observable environment by providing the agent with local sensations of the gradient areas surrounding its position. The local sensation was divided into four quadrants (Figure 1) to allow goal-directed-like navigation. By observing differences between the upper horizontal quadrants (q_i^0 and q_i^1), H_i^{sign} controlled the agent's orientation. AD_i^{sign} implements an appetitive-aversive behavioral axis (gradient ascent/descent) by comparing agent (*aGL*) and resource gradient location (*rGL*); understanding *aGL* as the mean gradient value between the four quadrants and *rGL* as the gradient value at its peak.

$$H_i^{sign} = \begin{cases} 1 & if\ q_i^0 < q_i^1 - th \\ -1 & if\ q_i^0 > q_i^1 + th \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$AD_i^{sign} = \begin{cases} 1 & if\ aGL_i^0 < rGL_i^1 - th \\ -1 & if\ aGL_i^0 > rGL_i^1 + th \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Internally-driven navigation

Outputs of the excitatory populations of the neural mass allostatic model are integrated with H_i^{sign} and AD_i^{sign} orientational signals to result in internally-driven goal-oriented-like navigation, following:

$$Navigation = c + \left(\sum_{i=1}^{nGrad} H_i^{sign} \cdot AD_i^{sign} \cdot Exc_i^{out\,put} \right) \cdot \frac{1}{nNeeds} \tag{7}$$

where c is a constant ensuring a default action going forward, $Exc_i^{out\,put}$ accounts for the output of the excitatory population i of the neural mass allostatic model, and $\frac{1}{nNeeds}$ is a normalization factor given the number of implemented needs.

Experimental design

A first study was conducted to evaluate the competence of the neural mass allostatic model in replicating rodent behavior during an open field test. In this study, a synthetic agent navigated a static simulated environment to defend its internal states of security and arousal. 50 experiments were carried out to analyze navigational patterns consistency and the internal dynamics of the agent along the simulation.

In the second study, the synthetic agent endowed with the neural mass allostatic model incorporated two distinct internal needs: thermoregulation and hydration. A dynamic simulated environment allowed us to assess the agent's ability to adapt its navigation according to environmental opportunities. 50 experiments were carried out to analyze the consistency of the navigational patterns, the internal dynamics taking place along the simulation, and the relationship between these internal dynamics and environmental changes.

In both the first and the second studies, the agent must constantly decide what internal drives should base its navigation on to maximize stability. This continuous decision-making condition represents a major difference from the original work in which the model's design is based on (Amil and Verschure, 2021) and is a novel challenge to overcome. Therefore, a third study was conducted to evaluate the advantages of inducting subcritical dynamics after need resolution. As in the second study, a synthetic agent navigated a dynamic environment to defend its internal states of hydration and thermoregulation. 50 experiments were conducted to assess the advantages of applying global inhibition after drive-completion behavior compared with the second study, where we did not consider such an effect.

Arenas

The arenas simulated in static and dynamic conditions were designed as two-dimensional 200 × 200 matrices incorporating one gradient per internal need. These gradients represented the location in the two-dimensional space where each internal need can be satisfied.

In the first study, we built a static environment following the gradient design of the open field test used in (Sanchez-Fibla et al., 2010). Thus, we implemented two gradients mapping the opportunities to calm security and arousal drives. The security gradient was designed as a Gaussian gradient with its peak in the top-left corner, representing the home base location. Meanwhile, the arousal gradient was designed with its peak in the center of the arena, representing the animal's exposure level. The configuration of these gradients aimed to replicate rodent navigational patterns, understood as a preference to occupy the home base and to explore the arena close to the walls (i.e., thigmotaxis) with occasional transversals (Figure 2A).

In the second and third studies, a dynamic environment implemented two gradients to map the opportunities to satisfy thermoregulation and hydration internal needs. We used a similar static gradient to security in the first study to

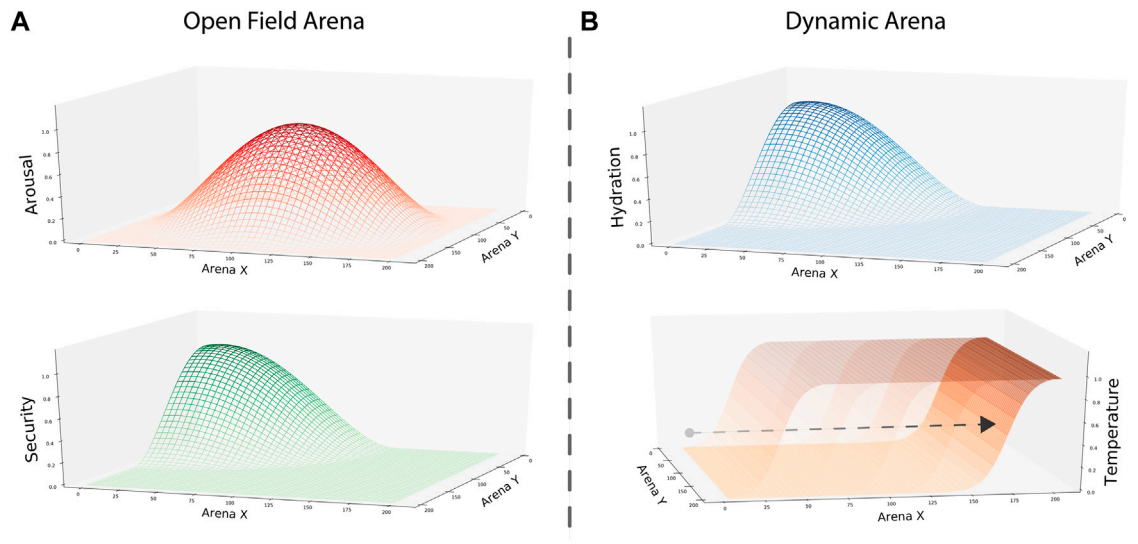


FIGURE 2
Gradients representing environmental opportunities to satisfy internal needs in each condition. **(A)** Arousal and security gradients were designed to replicate rodent behavior in an open field test. To represent the maximal level of exposure when exploring the center of the arena, we set the peak of the gradient in that location. Similarly, we used the peak of the security gradient to represent the home base location in one of the corners. **(B)** Hydration and temperature gradients were designed to test the performance of our model in a dynamic setting. Here, the hydration gradient was static, with its peak in one of the corners. In contrast, the temperature gradient changed over time. To do so, we built the gradient as a two-dimensional sigmoidal function where its *x*-intercept increased as the simulation evolved. Thus, the peak area where internal temperature increases shrunk, and the intermediate area between gradients increased over time.

represent the thirst-calming resource—nonetheless, the gradient representing the area where internal temperature increases changed over time. Designed as a two-dimensional sigmoidal function, the area providing temperature was large at the beginning of the simulation. However, as the simulation evolved, the *x*-intercept of our sigmoidal gradient increased, shortening the temperature gradient's peak area and establishing a larger gap between the two resource locations (Figure 2B). Thus, at the experiment's end, the temperature gradient's peak area only covers a short area of the arena at its bottom limit.

Efficiency, fairness and stability

To further understand the relationship between internal and external variables, in the second and third studies, we compute three measures traditionally used in game theory (Binmore, 2005; Hawkins and Goldstone, 2016; Freire et al., 2020). Efficiency provides a measure of how good the system is in maintaining the internal states close to their setpoints. Fairness designates whether there is any bias in the engagement of drives. Lastly, stability, computed as the mean square homeostatic error, comprises both the magnitude and the difference between homeostatic errors to provide a general measure of the agent performance.

To inform about the evolution of these measures throughout the experiment, they were calculated for every 1/100 faction of the duration of the experiment. Hence, mean internal and desired states only represented those elements in the considered fraction across the 50 experiments.

$$Efficiency = \frac{meanT + meanH}{nNeeds} \tag{8}$$

$$Fairness = \frac{|meanT - meanH|}{nNeeds} \tag{9}$$

$$Stability = 1 - \sum_{i=1}^{nGrad} (meanIS_i - meanDS_i)^2 \tag{10}$$

Results

Open field test

In the first study, as previous works did (Sanchez-Fibla et al., 2010), we aim to reproduce the navigational patterns of a rodent in an open field test. As in rodents, the trajectories of our agent showed preferences toward the walls (thigmotaxis) and the top-left corner (home base). At the same time, occasional transversals explored the center of the arena (Figure 3A, Supplementary Figure S5). The observed trajectory patterns resulted from the competition between two drives with different temporal

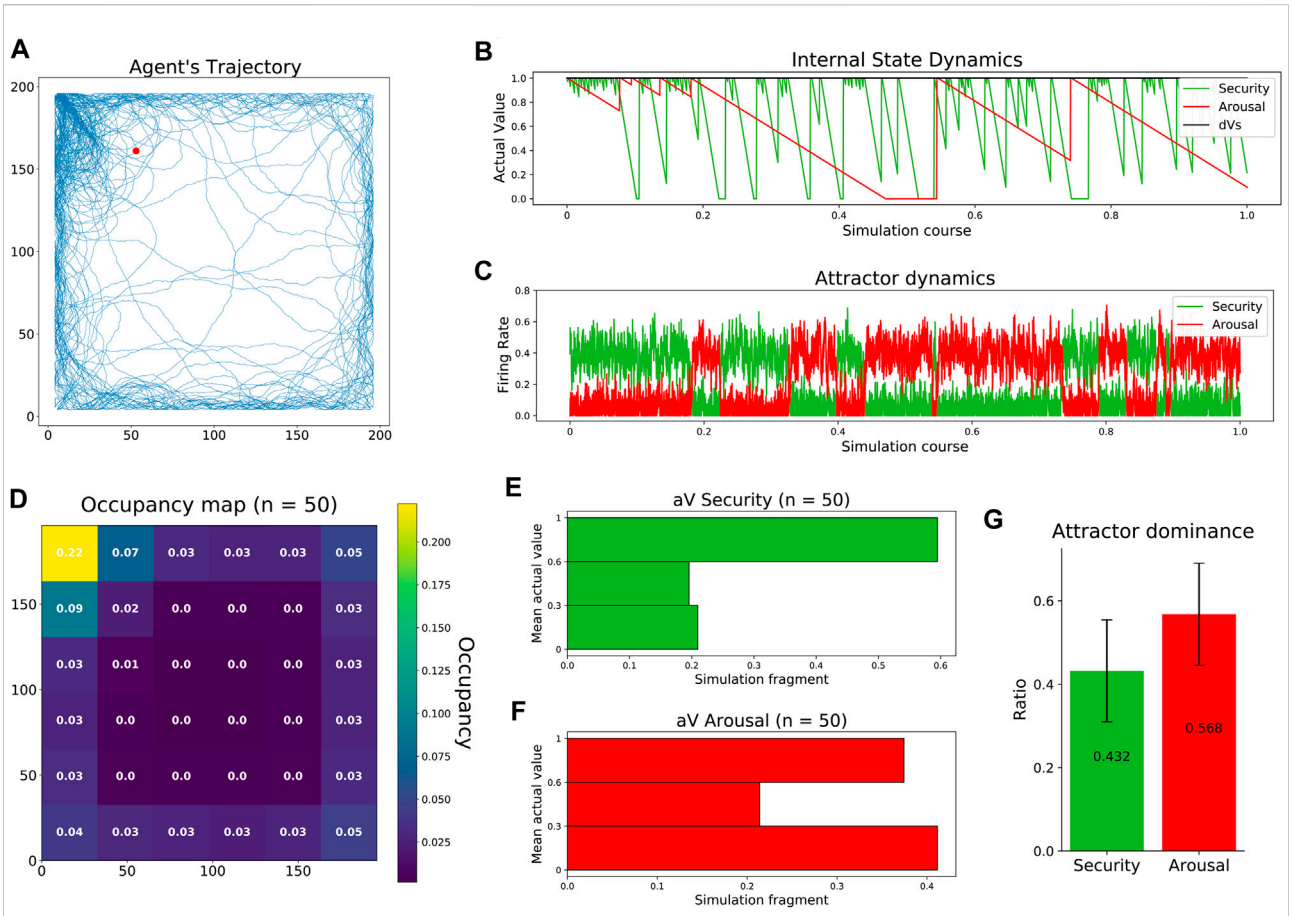


FIGURE 3
Replication of rodent behavior in an open field test. **(A)** Agent's trajectory tracked along a complete experiment. The red dot represents starting location, randomized across experiments. **(B)** Agent's internal state dynamics during a complete experiment (same as a). **(C)** Firing rates of the excitatory populations configuring the neural mass allostatic model during a complete experiment (same as a). **(D)** Occupancy map across 50 experiments with normalized values. **(E)** Distribution of security states across 50 experiments. **(F)** Distribution of arousal states across 50 experiments. **(G)** Mean attractor dominance across 50 experiments.

dynamics. Fast-decaying security made the agent constantly revisit its home base, while slow-decaying arousal allowed the agent to occasionally visit the center of the arena (Figure 3B). Competing dynamics between the two internal needs emerged when feeding the excitatory populations of our neural mass model with the corresponding homeostatic errors, allowing the dominant attractor to inhibit its opponent, thus resolving drive orchestration (Figure 3C). After carrying out 50 experiments, agent trajectories showed consistency in their occupancy pattern: a clear preference to visit the home base, navigate close to the walls, and avoid the center of the arena (Figure 3D). The distribution of the internal states during those 50 simulations was very informative. On the one hand, the agent maintained a high level of security during a large part of the simulations (Figure 3E). On the other hand, the state of arousal followed a bimodal distribution indicating the agent was either high or low aroused during the experiment. This bimodal distribution can be interpreted as follows: First, while the agent pursues security, an aversion toward the center of the arena induces low states of arousal. Then, low internal states, in turn, promote transversals that fully replenish the arousal of the agent (Figure 3F). Finally, attractor dominances, i.e., the simulation period where the firing rate of one excitatory neural pool exceeded the firing rate of the other, showed a balanced activation of both attractors with a slight bias toward arousal (Figure 3G).

Dynamic environment

In the second study, we assess the robustness of our novel neural mass allostatic model in a dynamic environment. In this condition, we observe that the agent's trajectory accurately tracks the environmental gradients (Figure 2B) occupying their peak areas

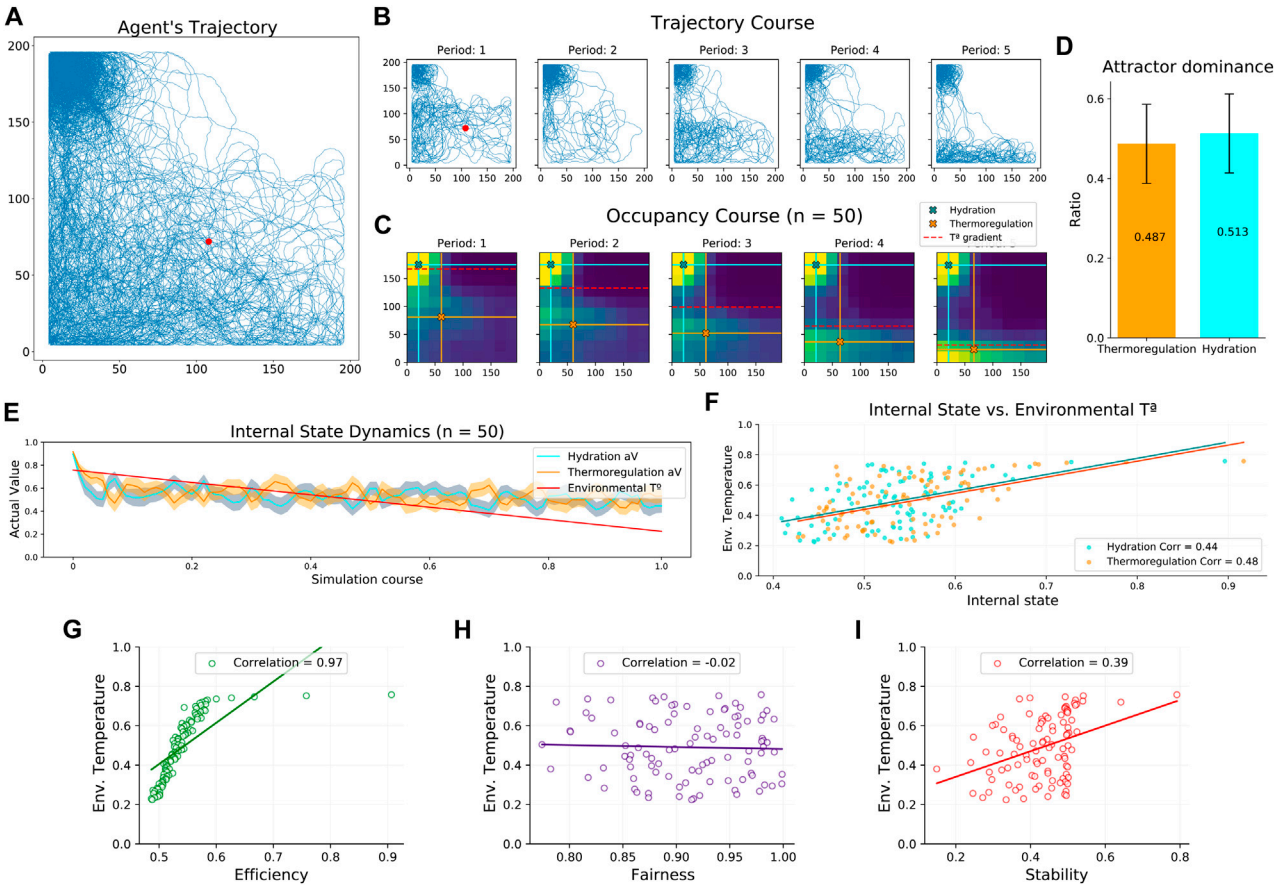


FIGURE 4
Agent's performance in a dynamic environment. **(A)** Agent's trajectory tracked along a complete experiment. The red dot represents starting location, randomized across experiments. **(B)** Agent's trajectory tracked along a complete experiment (same as a) divided into five periods. **(C)** Map of occupancy divided into five periods. Color crosses illustrate the mean position during hydration (blue) and thermoregulation (orange) attractor dominance. The red dashed line illustrates the variable slope location of the dynamic temperature gradient. **(D)** Mean attractor dominance **(E)** Mean agent internal state dynamics and environmental temperature. The shaded area indicates the internal state variance. **(F)** Mean internal state dynamics correlated with the environmental temperature. **(G)** Efficiency dynamics correlated with the environmental temperature. **(H)** Fairness dynamics correlated with the environmental temperature. **(I)** Stability dynamics correlated with the environmental temperature. Aggregated results from 50 experiments.

(Figure 4A, Supplementary Figure S6). Specifically, when dividing the complete experiment into five periods of equal length, we can observe that the agent's trajectory constantly adapted to the changes in the temperature gradient (Figure 4B). Thus, when initializing the experiment (period 1), the peak area of the temperature gradient covers a large part of the arena, and the agent navigates more extensively. However, in the last period (period 5), the peak area is reduced to a thin region close to the bottom border, and the agent's navigation adapts to it. This trajectory pattern was consistent across the 50 experiments, as the occupancy maps suggest (Figure 4C). Spearman correlation analysis indicated that the mean Y axis position of the agent during thermoregulation highly correlated with the temperature gradient's slope location ($r(498) = 0.96$, $p < 0.001$), confirming the agent's trajectory adaptation. Once again, attractor dominance was balanced (Figure 4D), suggesting that although the environment asymmetrically reduces the opportunities to fulfill the agent's internal needs, the neural mass model imposes a well-balanced competition without neglecting any of the drives. Indeed, the mean internal state across the experiment shows that thermoregulation and hydration levels are well maintained without biases (Figure 4E), and both are equally correlated with the environmental temperature (mean value of the temperature gradient) (Figure 4F). To understand in detail the relationship between internal states and environmental dynamics, we studied this relationship in terms of efficiency, fairness, and stability metrics. Correlating these measures with the environmental temperature, we observed that efficiency strongly depends on the opportunities to fulfill internal needs

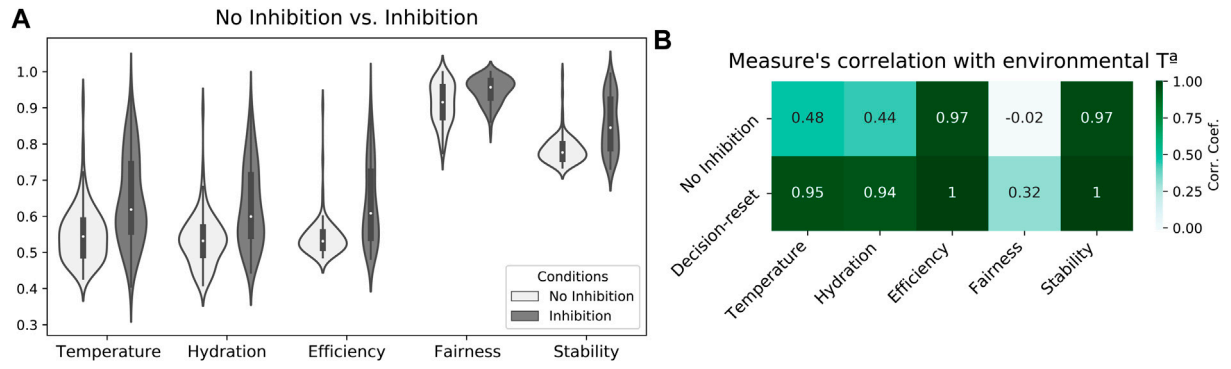


FIGURE 5
Decision-reset supports the agent's internal stability. **(A)** Comparison of temperature and hydration internal states and efficiency, fairness, and stability scores between no-inhibition (study 2) and decision-reset (study 3) conditions. **(B)** Correlation between internal states and game theory measures with environmental temperature for both no inhibition and decision-reset conditions.

(Figure 4G). This result was expected, given the environmental dynamics. When the environmental temperature decreases (i.e., temperature gradient peak shrinks), gradients' peak areas are more distant, forcing the agent to navigate a larger area where its internal states decrease. However, environmental changes do not affect the fairness level (Figure 4H), validating that the balance of attractor dominance is defended even when receiving asymmetric influences from the environment. Lastly, stability reports difficulties minimizing the mean square homeostatic error as temperature decreases (Figure 4I). According to our previous analysis, this result would be better explained by environmental influences on efficiency than fairness.

Criticality-driven decision-reset after decision accomplishment

Finally, a third study explores the potential limitations of attractor-based allostatic models. Specifically, by inhibiting the model's excitatory neural populations (once drive-completion behaviors have been performed), we explore whether sustained attractor forces could hamper individual self-regulation in study 2. Results showed that agents widely benefit from inducing a critical regime to set the initial conditions for each cycle of allostatic orchestration. Specifically, mean internal states along the simulation (calculated every 1/100 fraction of the experiment) across 50 simulations indicated that agents maintained their internal states better when applying decision-reset (Inhibition). (Figure 5A). Consequently, increased internal states also resulted in increased efficiency, fairness, and stability scores. A Mann-Whitney test indicated statistically significant differences when comparing the two conditions in each measure, $U(N_{No-inhibition} = 100, N_{inhibition} = 100)$, $p < 0.001$). Furthermore, this enhanced performance occurred in larger alignment with the environmental temperature (Figure 5B), which indicates the agent's internal states decreased when the scarcity of resources prevented a better self-regulatory strategy. Altogether, these results suggest that attractor forces sustained after drive-completion behaviors hampered allostatic orchestration, and criticality-driven decision-reset provides an effective mechanism to facilitate a more flexible decision-making process.

Discussion

Previous computational works have contributed to explaining animal self-regulatory behavior (Sanchez-Fibla et al., 2010; Keramati and Gutkin, 2014; Jimenez-Rodriguez et al., 2020). However, these models have not been validated in a broad range of tasks with emphasis on their robustness in the face of varying task conditions, while in parallel, their grounding in the brain mechanisms underlying allostatic orchestration is not fully elucidated. We hypothesize that attractor dynamics originate from the competing relation between different hypothalamic interoceptive nuclei and their loops through the zona incerta, which are further perturbed by the superior colliculus (input and orienting control), and the central gray for triggering reactive species-specific behaviors. To test the adequacy of allostatic attractor dynamics in orchestrating internal needs, we built a biologically-constrained computational model of allostatic orchestration. The resultant model implements competition between internal drives by comprising two drive-sensitive excitatory neural populations that apply mutual and feedback inhibition. To allow a synthetic agent to navigate a virtual external environment based on its internal state, we endowed it with the ability to orientate and perform basic appetitive-aversive navigational behaviors. In this manner, the hypothalamus, superior colliculus, and zona incerta-periaqueductal gray axis are represented in our model, providing a first approximation of the core behavior system (Merker, 2013; Verschure, 2016).

Deploying the allocentric synthetic agent in a virtual environment, we tested the competencies and robustness of the neural mass allostatic model. Specifically, our model allowed the agent to 1) navigate an open field reproducing rodent behavior, 2) adapt its navigation to environmental changes without neglecting any internal needs, and 3) benefit from criticality reset to optimize the interoceptive-driven decision-making process. Altogether, our results supported our hypothesis empirically validating attractor dynamics as an inherent feature of the hypothalamic circuitry that can underlie robust allostatic orchestration. However, the attractor dynamics that emerge in our model need to be further validated by physiological data, ideally by direct neural recordings of the hypothalamic interoceptive nuclei. Lacking ground truth benchmarking data, the temporal dynamics of the homeostatic markers employed in this work were arbitrarily set, representing an additional methodological limitation. Moreover, now that we have assessed the robustness and fidelity of the model to track environmental dynamics, we need to assess its scalability.

Additionally, our work opens new questions in the quest to understand self-regulation. For instance, a better understanding of the interplay between stress and risky internal states is needed. In our studies, the level of competition between neural populations remained fixed at an optimal point. However, the ratio between mutual and shared inhibition could be modulated by stress markers such as acetylcholine (Kawaguchi, 1997). How stress can modulate the competition held between internal drives is a question that remains unanswered. Another open question is how allostatic orchestration interplays with predictive and contextual allostasis to assist each other. We propose to structure allostasis as a cognitive architecture following the four organizational layers of the Distributed Adaptive Control theory of mind (Verschure, 2012). Here, a first somatic level endows the agent with predefined internal needs. Then, the reactive level rapidly orchestrates behaviors directed to calm the most salient drives. In parallel, the adaptive layer learns associations between external and internal events to predict future homeostatic errors and solve them in anticipation. Lastly, a contextual layer enriches the representation of the self-regulatory strategies and enables their memorization.

Finally, contributions to the field of allostasis can expand the boundaries of self-regulation in its understanding and implementation. On the one hand, drives and motivations should also respond to a hierarchical organization that encodes not only priority but also abstraction, virtualization, and replaceability of the drive (Verschure, 2016). Maslow's hierarchy was a first approximation to capture this organization (Maslow, 1943); however, 7 decades later, this theory has not been further advanced in the face of new insights into the dynamics of self-regulation. How different needs within the same priority level organize or how new needs arise and substitute previous ones are questions not answered yet. On the other hand, further investigation could shed light on how a hierarchical organization of needs fits with a neural mass model. Secondly, general-purpose robot instantiations can benefit greatly from the self-regulatory principles described here. Previous works elaborated on how a multi-agent robotic recycling plant can implement homeostatic and allostatic principles to self-organize at both single-agent and large-scale levels (Rosado and Verschure, 2020). Similar architectures can be designed and implemented to advance a variety of scenarios where robot autonomy is key such as robot delivery or space exploration. Indeed, autonomy in artificial intelligence and robotics places the question of self-regulation at the center of the research of synthetic embodied cognition and consciousness.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

OGR, AFA, and ITF designed the protocol. OGR conceived and conducted the experiments, analyzed the results, and wrote the manuscript. PV initiated and supervised the research. All authors reviewed and approved the manuscript.

Funding

This study was supported by Counterfactual Assessment and Valuation for Awareness Architecture—CAVAA (European Commission, EIC 101071178) and Hybrid Human-Robot RECYcling plant for electriCal and eLEctRonic equipment—HR-RECYCLER (Horizon 2020, project ID: 820742).

Acknowledgments

We thank our colleagues from the Synthetic, Perceptive, Emotive and Cognitive Systems Lab (SPECS) who provided valuable feedback during the entire conception of the present work. We would also like to show our gratitude to those students from the Cognitive Systems and Interactive Media master (University Pompeu Fabra, Barcelona) who showed interest in the research line and contributed indirectly with parallel research projects. Last but not the least, we are also grateful to the reviewers for taking the time and effort to review the manuscript and provide such valuable comments and suggestions.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2022.1052998/full#supplementary-material

References

Amil, A. F., and Verschure, P. F. (2021). Supercritical dynamics at the edge-of-chaos underlies optimal decision-making. *J. Phys. Complex.* 2, 045017. doi:10.1088/2632-072x/ac3ad2

Basu, R., Gebauer, R., Herfurth, T., Kolb, S., Golipour, Z., Tchumatchenko, T., et al. (2021). The orbitofrontal cortex maps future navigational goals. *Nature* 599, 449–452. doi:10.1038/s41586-021-04042-9

Bernard, C. (1865). Introduction à l'étude de la médecine expérimentale. Paris, France: J. B. Baillière et fils.

Binmore, K. (2005). *Natural justice*. Oxford: Oxford University Press.

Blouet, C., and Schwartz, G. J. (2010). Hypothalamic nutrient sensing in the control of energy homeostasis. *Behav. Brain Res.* 209, 1–12. doi:10.1016/j.bbr.2009.12.024

Burnett, C. J., Li, C., Webber, E., Tsaousidou, E., Xue, S. Y., Brüning, J. C., et al. (2016). Hunger-driven motivational state competition. *Neuron* 92, 187–201. doi:10.1016/j.neuron.2016.08.032

Cabanac, M. (1971). Physiological role of pleasure: a stimulus can feel pleasant or unpleasant depending upon its usefulness as determined by internal signals. *Science* 173, 1103–1107. doi:10.1126/science.173.4002.1103

Cannon, W. B. (1939). *The wisdom of the body*. New York, NY: W. W. Norton and company INC.

Freire, I. T., Moulin-Frier, C., Sanchez-Fibla, M., Arsiwalla, X. D., and Verschure, P. F. (2020). Modeling the formation of social conventions from embodied real-time interactions. *PloS one* 15, e0234434. doi:10.1371/journal.pone.0234434

Gould, T. D., Dao, D. T., and Kovacsics, C. E. (2009). *The open field test. Mood and anxiety related phenotypes in mice*, 1–20. Berlin, Germany: Springer Science+Business Media.

Hawkins, R. X., and Goldstone, R. L. (2016). The formation of social conventions in real-time environments. *PloS one* 11, e0151670. doi:10.1371/journal.pone.0151670

Jimenez-Rodriguez, A., Prescott, T. J., Schmidt, R., and Wilson, S. (2020). "A framework for resolving motivational conflict via attractor dynamics," in *Conference on biomimetic and biohybrid systems* (Germany: Springer), 192–203.

Kawaguchi, Y. (1997). Selective cholinergic modulation of cortical gabaergic cell subtypes. *J. neurophysiology* 78, 1743–1747. doi:10.1152/jn.1997.78.3.1743

Keramati, M., and Gutkin, B. (2014). Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife* 3. doi:10.7554/elife.04811

Laurençon, H., Ségerie, C.-R., Lussange, J., and Gutkin, B. S. (2021). Continuous homeostatic reinforcement learning for self-regulated autonomous agents. *arXiv preprint arXiv:2109.06580*.

Lee, C. R., Chen, A., and Tye, K. M. (2021). The neural circuitry of social homeostasis: Consequences of acute versus chronic social isolation. *Cell* 184, 1500–1516. doi:10.1016/j.cell.2021.02.028

Li, M., Han, Y., Aburn, M. J., Breakspear, M., Poldrack, R. A., Shine, J. M., et al. (2019). Transitions in information processing dynamics at the whole-brain network level are driven by alterations in neural gain. *PLoS Comput. Biol.* 15, e1006957. doi:10.1371/journal.pcbi.1006957

Ma, Z., Turrigiano, G. G., Wessel, R., and Hengen, K. B. (2019). Cortical circuit dynamics are homeostatically tuned to criticality *in vivo*. *Neuron* 104, 655–664.e4. doi:10.1016/j.neuron.2019.08.031

Marshall, J. A., Favreau-Peigné, A., Fromhage, L., Mcnamara, J. M., Meah, L. F., and Houston, A. I. (2015). Cross inhibition improves activity selection when switching incurs time costs. *Curr. Zool.* 61, 242–250. doi:10.1093/czoolo/61.2.242

Maslow, A. (1943). *A theory of human motivation*. psychological review no 50. Washington, DC: American Psychological Association.

Maturana, H. R., and Varela, F. J. (1991). *Autopoiesis and cognition: The realization of the living*, 42. Germany: Springer Science & Business Media.

Merker, B. (2013). The efference cascade, consciousness, and its self: Naturalizing the first person pivot of action control. *Front. Psychol.* 4, 501. doi:10.3389/fpsyg.2013.00501

Nakamura, K. (2011). Central circuitries for body temperature regulation and fever. *Am. J. Physiology-Regulatory, Integr. Comp. Physiology* 301, R1207–R1228. doi:10.1152/ajpregu.00109.2011

Osterhout, J. A., Kapoor, V., Eichhorn, S. W., Vaughn, E., Moore, J. D., Liu, D., et al. (2022). A preoptic neuronal population controls fever and appetite during sickness. *Nature* 1–8, 937–944. doi:10.1038/s41586-022-04793-z

Papies, E. K., and Aarts, H. (2016). Automatic self-regulation: From habit to goal pursuit. *Handbook of self regulation: Research, theory, and applications* (New York: Guilford Press).

Pavlov, I. P. (1955). *Selected works*, 1. Russia: Foreign Languages Publishing House.

Qian, S., Yan, S., Pang, R., Zhang, J., Liu, K., Shi, Z., et al. (2022). A temperature-regulated circuit for feeding behavior. *Nat. Commun.* 13, 4229–4317. doi:10.1038/s41467-022-31917-w

Rosado, O. G., and Verschure, P. F. (2020). "Distributed adaptive control: An ideal cognitive architecture candidate for managing a robotic recycling plant," in *Conference on biomimetic and biohybrid systems* (Germany: Springer), 153–164.

Sanchez-Fibla, M., Bernardet, U., Wasserman, E., Pelc, T., Mintz, M., Jackson, J. C., et al. (2010). Allostatic control for robot behavior regulation: a comparative rodent-robot study. *Adv. Complex Syst.* 13, 377–403. doi:10.1142/s0219525910002621

Sherrington, C. S. (1906). *The integrative action of the nervous system*, 35. New Haven: Yale University Press.

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiology Behav.* 106, 5–15. doi:10.1016/j.physbeh.2011.06.004

Sterling, P. (1988). "Allostasis: a new paradigm to explain arousal pathology," in *Handbook of life stress, cognition and health*. Hoboken, New Jersey: John Wiley and Sons.

Sterling, P. (2020). *What is health?: Allostasis and the evolution of human design*. Cambridge: MIT Press.

Strecker, R., Nalwalk, J., Dauphin, L., Thakkar, M., Chen, Y., Ramesh, V., et al. (2002). Extracellular histamine levels in the feline preoptic/anterior hypothalamic area during natural sleep–wakefulness and prolonged wakefulness: an *in vivo* microdialysis study. *Neuroscience* 113, 663–670. doi:10.1016/s0306-4522(02)00158-6

Tschantz, A., Barca, L., Maisto, D., Buckley, C. L., Seth, A. K., and Pezzulo, G. (2022). Simulating homeostatic, allostatic and goal-directed forms of interoceptive control using active inference. *Biol. Psychol.* 169, 108266. doi:10.1016/j.biopsycho.2022.108266

Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592. doi:10.1037/0033-295x.108.3.550

Verschure, P. F. (2012). Distributed adaptive control: a theory of the mind, brain, body nexus. *Biol. Inspired Cogn. Archit.* 1, 55–72. doi:10.1016/j.bica.2012.04.005

Verschure, P. F. (2016). Synthetic consciousness: the distributed adaptive control perspective. *Phil. Trans. R. Soc. B* 371, 20150448. doi:10.1098/rstb.2015.0448

Wiener, N. (1948). *Cybernetics*. New York: John Wiley & Sons.

Yuan, Y., Wu, W., Chen, M., Cai, F., Fan, C., Shen, W., et al. (2019). Reward inhibits paraventricular crh neurons to relieve stress. *Curr. Biol.* 29, 1243–1251.e4. doi:10.1016/j.cub.2019.02.048

Zimmerman, C. A., Leib, D. E., and Knight, Z. A. (2017). Neural circuits underlying thirst and fluid homeostasis. *Nat. Rev. Neurosci.* 18, 459–469. doi:10.1038/nrn.2017.71

frontiers | Frontiers in Robotics and AI

TYPE Perspective
PUBLISHED 06 February 2023
DOI 10.3389/frobt.2023.1030416

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Stefania Costantini,
University of L'Aquila, Italy

*CORRESPONDENCE
Nicolas Coucke,
nicolas.coucke@ulb.be
Mary Katherine Heinrich,
mary.katherine.heinrich@ulb.be

†These authors have contributed equally to this work and share first authorship.

SPECIALTY SECTION
This article was submitted to Computational Intelligence in Robotics, a section of the journal Frontiers in Robotics and AI

RECEIVED 28 August 2022
ACCEPTED 23 January 2023
PUBLISHED 06 February 2023

CITATION
Coucke N, Heinrich MK, Cleeremans A and Dorigo M (2023), Learning from humans to build social cognition among robots. *Front. Robot. AI* 10:1030416. doi: 10.3389/frobt.2023.1030416

Learning from humans to build social cognition among robots

Nicolas Coucke[1,2*†], Mary Katherine Heinrich[1*†], Axel Cleeremans[2] and Marco Dorigo[1]

[1]IRIDIA, Université Libre de Bruxelles, Brussels, Belgium, [2]Consciousness, Cognition and Computation Group, Center for Research in Cognition and Neurosciences, Université Libre de Bruxelles, Brussels, Belgium

Self-organized groups of robots have generally coordinated their behaviors using quite simple social interactions. Although simple interactions are sufficient for some group behaviors, future research needs to investigate more elaborate forms of coordination, such as social cognition, to progress towards real deployments. In this perspective, we define social cognition among robots as the combination of social inference, social learning, social influence, and knowledge transfer, and propose that these abilities can be established in robots by building underlying mechanisms based on behaviors observed in humans. We review key social processes observed in humans that could inspire valuable capabilities in robots and propose that relevant insights from human social cognition can be obtained by studying human-controlled avatars in virtual environments that have the correct balance of embodiment and constraints. Such environments need to allow participants to engage in embodied social behaviors, for instance through situatedness and bodily involvement, but, at the same time, need to artificially constrain humans to the operational conditions of robots, for instance in terms of perception and communication. We illustrate our proposed experimental method with example setups in a multi-user virtual environment.

KEYWORDS

artificial social cognition, embodied cognition, self-organization, robot swarms, multi-robot systems, artificial intelligence, artificial general intelligence, social robots

Introduction

AI research has greatly advanced, but when interaction with other agents is required, existing algorithms easily break down (Bard et al., 2020). Social interaction and social embodiment are still underexplored in artificial general intelligence (Bolotta and Dumas, 2022) and in groups of intelligent robots. While there is some robotics research on social cognition, it focuses on human-robot interaction (Henschel et al., 2020), e.g., how a robot interprets the intentions of a human, not on interactions *among* robots. It is important to note that what looks like social cognition is not necessarily social cognition. For instance, agents or robot controllers made by reinforcement learning might behave in ways that look socially cognizant in some situations, but this might only be appearance—i.e., the underlying behavioral phenomena are not there—so the illusion will break down when exposed to more situations.

Robots can coordinate with each other by using, e.g., centralized control or self-organization. In multi-robot systems that are not self-organized, robots are directed to follow a centrally coordinated plan using explicit commands or global references. In this paper, we are interested exclusively in robot groups that include aspects of self-organization, because social cognition depends on some degree of individual autonomy. If a robot is essentially a remote-controlled sensor or actuator, it does not engage in social cognition.

In existing research on self-organized robot groups, the individuals are usually quite simple and often rely on indiscriminate, naïve interactions. Indeed, swarm robotics research has shown that no advanced cognition or elaborate social negotiation is needed to self-organize certain group behaviors (e.g., Nouyan et al., 2009; Rubenstein et al., 2014; Valentini et al., 2016). However, it has been argued that there are still significant gaps for robot swarms to be deployment-ready, and that the future of swarm robotics research should concentrate on more elaborate forms of self-organized coordination (Dorigo et al., 2020; 2021), such as self-organized hierarchy (Mathews et al., 2017; Zhu et al., 2020) or behavioral heterogeneity (Kengyel et al., 2015).

In this perspective, we argue that another important direction for future study should be social cognition. Robot groups successfully equipped with social cognition could engage in elaborate coordination without sending each other large amounts of data. Some aspects of robot behavior could be mutually predictable, for instance by robots maintaining good internal models of each other. Socially cognitive robots could have improved group performance, e.g., by not destructively interfering with each other (which requires time and effort to resolve) and not accidentally disrupting each other's sub-goals while attempting to reach a common goal.

In cognitive robotics, research on individual robots such as humanoids is very advanced (Cangelosi and Asada, 2022), even on each of the six key attributes of artificial cognitive systems (Vernon, 2014): action, perception, autonomy, adaptation, learning, and anticipation. Comparatively, cognition in swarm robotics is still in its beginning stages. While cognitive robot swarms can be autonomously capable of collective action, perception, and in some cases adaptation (Heinrich et al., 2022), we do not yet know how to make robot swarms that can autonomously learn and anticipate as a collective, in such a way that the group behavior is greater than the sum of its parts. We propose that studying social cognition could help us advance the autonomous collective capabilities of groups of robots.

Socially cognitive robots: Our perspective

Our perspective is summarized as follows: social cognition among robots can be built by developing artificial social reasoning capabilities based on behaviors observed in humans.

Frith (2008) has defined social cognition in humans as "the various psychological processes that enable individuals to take advantage of being part of a social group" and Frith and Frith (2012) have further specified that a substantial portion of these psychological processes are for learning about and making predictions about other members of the social group. The mechanisms of social cognition in humans include social signalling, social referencing, mentalizing (i.e., tracking of others' mental states, intended actions, objectives, and opinions), observational learning (e.g., social reward learning, mirroring), deliberate knowledge transfer (e.g., teaching), and sharing of experiences through reflective discussion (Frith, 2008; Frith and Frith, 2012). Crucially, social cognition is also defined as "not reducible to the workings of individual cognitive mechanisms" (De Jaegher et al., 2010).

Although some social abilities such as simple social interaction are well-developed among robots, most of the abilities contained in Frith (2008)'s definition of social cognition are lacking, and could provide significant performance benefits. For instance, the transfer of information between robots is well understood, but much less so the transfer of knowledge, especially implicitly:

We define social cognition among robots as the following set of abilities:

1. **Social inference**—inferring the opinions, intended next actions, and overall goals of other robots in the same social group, using interpretation of social signals;
2. **Social learning**—learning information about which actions to adopt or avoid based on observations of each other's behaviors and social signalling;
3. **Social influence**—deliberately influencing each other's (socially inferred) internal states using social signaling; and
4. **Knowledge transfer**—transferring high-level knowledge using social interaction, e.g., using implicit demonstration or explicit instruction.

Currently, robots are well-equipped with some of the requirements for these abilities, such as simple social interactions, but lack other crucial requirements such as explicit social reasoning. Although research has shown that no social cognition is needed for simple group behaviors in robots, it is an open challenge how to accomplish more advanced behaviors in a fully self-organized way. Some of the significant unresolved technical challenges for advanced self-organization among robots, which we believe social cognitive abilities could contribute to, are the following:

- autonomously anticipating which actions should be taken in an environment filled with other autonomous robots,
- collectively defining an explicit goal that was not pre-programmed and collectively directing the robot group towards it,
- making online inferences about other robots' current states and future behaviors, and adapting their coordination strategies accordingly, even while moving at high speed in dynamic unknown environments, and
- designing self-organization among robots such that the resulting group behaviors, although not completely predictable, are safe and trustable.

We propose that socially cognitive robots can in part be developed by learning from the social cognition processes of humans in certain experimental conditions. In order to have the potential to transfer observed behaviors and capabilities from humans to robots, we believe experiments with human subjects must be conducted in a platform that allows experimental setups to be: on one hand, realistic enough to study **embodied** human behavior, but on the other hand, **constrained** and simplified enough to approximate the operational conditions of robots.

State of the art

Artificial social learning and artificial mentalizing

Many examples of artificial learning exist that seem relevant to the mechanisms of social cognition. However, key social aspects are not present in these existing methods: for instance, reward learning has

been demonstrated in robots (e.g., Daniel et al., 2015) but learning of social rewards among robots has not been studied. Likewise, robots learning by interacting with and observing other robots has been demonstrated (e.g., Murata et al., 2015), but not for the learning of socially relevant information nor to build behaviors among robots that are irreducible to the knowledge held by robots individually.

Currently, the most advanced research towards artificial social cognition can be seen in multi-agent reinforcement learning. In basic approaches, each agent would use reinforcement learning individually, treating other agents as part of the environment. In more elaborate existing approaches, agents are trained to model each other and several types of artificial mentalizing have been demonstrated (Albrecht and Stone, 2018). For example, in the Deep Reinforcement Opponent Network (DRON), one agent learns the representation of the opponent's policy (He et al., 2016). In another example, an agent uses itself as the basis to predict another agent's actions (Raileanu et al., 2018). One approach using a "Theory of Mind" network has even produced agents that can explicitly report inferred mental states of other agents and pass the classic "false belief test" for understanding the mental states of others (Rabinowitz et al., 2018). Current efforts in multi-agent learning use cooperative games such as Hanabi as benchmarks, which involves inferring the mental states of others and using that information to collaborate (Bard et al., 2020). For the development of artificial social cognition, the next step for this line of research would be to situate the mentalizing behaviors within the full set of social cognition mechanisms, including social influence and social reward learning (*cf.* Olsson et al., 2020).

Social cognition transfer between humans and robots

Robots have been used as experimental tools for the study of embodied social cognition. For instance, a variety of devices have been used to automatically provide synthetic social stimuli to animals in a naturalistic way (Frohnwieser et al., 2016). Similarly, the effect of humanoid robots on human social cognition has been broadly studied (Wykowska et al., 2016). Social robots in the context of human-robot interaction have also been investigated (e.g., Dautenhahn, 2007). However, to the best of our knowledge, no studies have looked at expanding these robot use cases into embodied artificial social cognition among robots, and no work apart from our own has proposed using experiments with humans to contribute to building social cognition among robots.

Directions for future research

Advanced group capabilities seen in humans can inspire similar capabilities in robots. For example, the human capabilities of selecting and following leaders (Van Vugt, 2006) and re-organizing communication networks around individuals with better information (Almaatouq et al., 2020) have recently inspired the development of self-organized hierarchies for robots, for instance using physical (Mathews et al., 2017) or wireless connections (Zhu et al., 2020). In the following sections, we identify cognitive processes used by humans in social situations that would be valuable for robot groups, and propose them as future research directions for building social cognition among robots.

Social heuristics and action selection

Humans often use cognitive processes known as "heuristics" to select actions in social situations. In humans, heuristics are defined as action selection strategies that usually deviate from economic rationality or Bayesian optimality but which facilitate a rapid action selection when time and knowledge about a situation are limited (Hertwig and Herzog, 2009). The hidden states of other agents cannot be directly observed, so the outcome of a social situation always has a high degree of uncertainty—selecting the optimal action is computationally intractable (Seymour and Dolan, 2008).

In humans, heuristics can involve continuous integration of multiple variables or sources of information, for example when deciding on a walking direction based on the position of other walking individuals (Moussaid et al., 2011). In psychology and neuroscience, action selection is often characterized as the result of an accumulation process, in which evidence that supports a certain decision or action is accumulated over time (Ratcliff and McKoon, 2008). A certain action is taken when the accumulated evidence crosses some threshold. The sources and manner of evidence integration can be determined by social heuristics. For example, evidence accumulation frameworks can characterize how humans use a "follow the majority heuristic" during social decision making (Tump et al., 2020), as well as how humans base their own movements on those of others during embodied competitive interactions (Lokesh et al., 2022).

Coupling, alignment, and mirroring

Humans often mirror each other's behaviors and can participate in a "coupling" behavior through reciprocal interactions. Implicit coupling can occur between physiological states (for example, synchronization of heartbeats and breathing rhythms). Explicit sensorimotor coupling involves mutual prediction of each other's actions and facilitates coordinated action sequences (Dumas and Fairhurst, 2021). On a higher cognitive level, reciprocal interactions can create alignment between internal cognitive states, which in turn facilitates better mutual prediction of actions (Friston and Frith, 2015).

Humans can also disengage from social interactions and instead mirror (or "simulate") others' actions as a type of internalized action (Buzsáki, 2019, p. 131). This capacity is supported by the mirror neuron system, which is active when observing and when executing a movement (Rizzolatti and Craighero, 2004). Internal simulation aids in understanding others' intentions and in selecting complementary actions (Newman-Norlund et al., 2007).

Mentalizing and shared representations

Simply mirroring the mental states of others is often not sufficient to infer their opinions, objectives, or intended actions (Saxe, 2005). Therefore, coupling and mirroring are often complemented in humans by higher-level cognition about others' beliefs, desires, and intentions, taking into account factors such as context and memory (Sebanz et al., 2006). This requires mentalizing, a process of inference about others' changing mental states, beyond simple mirroring (Frith and Frith, 2012).

For example, mentalizing based on observations of others' gazes facilitates taking others' perspectives into account and tracking their

beliefs about a shared environment or world (Frith and Frith, 2012). By observing others' movements, humans can also infer the confidence that others have in their beliefs (Patel et al., 2012) and the intentions that underlie their actions (Baker et al., 2009). Crucially, humans also mentalize based on third-party observations of others' interactions, and then estimate the social relationships between them (Ullman et al., 2009).

Tracking others' goals and beliefs helps humans to distinguish which subset of their action representations are shared with others. Shared representations aid in predicting and interpreting the actions of others in the context of a joint goal, and in selecting complementary actions. For instance, by tracking others' beliefs, an individual can recognize when communication or signalling is needed to facilitate smooth coordination (Pezzulo and Dindo, 2011).

Outcome monitoring

Humans monitor behaviors and detect errors when taking actions directed towards a certain goal (Botvinick et al., 2001). If an individual recognizes another making what might be an error, in pursuit of a shared goal, the individual needs to then distinguish whether it was indeed an error, or whether their goals are misaligned.

Humans also monitor whether actions have their intended outcomes, as well as whether a certain action and certain outcome actually have a causal link. This results in a greater or lesser sense of agency over a certain action or outcome (Haggard and Chambon, 2012), which in turn impacts how an individual acts in social situations. Agency can be modulated in a variety of ways: joint agency when acting together with others, vicarious agency when influencing the actions of others, or violated agency when actions are interfered with by others (Silver et al., 2020). The modulated sense of agency in humans helps shape an individual's monitoring of links between actions, errors, and outcomes.

From humans to robots: An experimental method

Robots are embodied agents with specific morphologies and specific perception and action capabilities that differ from (and are often far more limited than) those of humans. To gain insights from human social cognition that are relevant to robots, human subjects would need to be studied in an experimental platform that: 1) allows them to engage in embodied social behaviors, but also 2) allows enough constraints to artificially expose humans to the operational conditions of robots. We propose that behavioral experiments conducted with humans controlling avatars in virtual environments can achieve this trade-off.

Balancing embodiment and constraints in virtual environments

Existing experiments on human social cognition have mostly been conducted in highly controlled single-person paradigms which lack embodiment. We identify the following five aspects of embodiment that we propose human-controlled avatars in new virtual environments will need, for the study of embodied human social cognition.

1. **Situatedness**: An agent takes actions while being part of a situation, rather than by observing the situation from the outside (Wilson, 2002).
2. **Sensory and action shaping**: By taking actions (e.g., moving their bodies) in the environment, agents can actively change the flow of their sensory inputs as well as the potential effects of their actions (Gordon et al., 2021).
3. **Bodily involvement**: The bodily state and/or morphology of the agent—as well as the agent's bodily relation to the bodies of other agents—can be involved in cognition (Wilson, 2002).
4. **Interaction cascades**: Agents can engage with each other in such a way that actions by one can influence reciprocal actions by another, resulting in cascades of interactions and behaviors (Dale et al., 2013).
5. **High bandwidth**: There can be high bandwidth of implicit or explicit information exchange between agents (Schilbach et al., 2013).

Complementarily, we identify the following constraints that will also need to be possible in the virtual environment.

1. **Body and action**: Human-controlled avatars can be equipped with morphology features and action capabilities that are similar to those of relevant robots.
2. **Perception**: When controlling an avatar, a human subject can be limited to sensory inputs similar to those of relevant robots (e.g., restricted visual information).
3. **Communication**: Human-controlled avatars can be limited to communication and signalling capabilities that are similar to those available to relevant robots.
4. **Hidden states**: Human subjects can be required to explicitly report information about hidden states (e.g., their current opinion or confidence level) that is not directly observable from their behavior but would be available to an experimenter if using relevant robots.

Unconstrained real-world social situations would fulfill all listed requirements for embodiment, but would lack control and interpretability. Virtual environments enable certain aspects of embodiment while at the same time ensuring a degree of control of the situation for the experimenter.

Example: Using the virtual environment HuGoS

To the best of our knowledge, no off-the-shelf virtual environment was available to meet these requirements, so we built a tool in Unity3D called "HuGoS: Humans Go Swarming" (Coucke et al., 2020; 2021) that we could use to study human behavior in embodied scenarios similar to those in which robots operate. To illustrate the features that we propose for a virtual environment for studying transferable social cognition, we describe two example experimental setups in HuGoS.

Collective decision making

Collective decision making has been widely studied in swarm robotics (Valentini et al., 2017), but many gaps still remain (Khaluf et al., 2019). Collective decisions have also been extensively studied in humans (Kameda et al., 2022), but not typically in

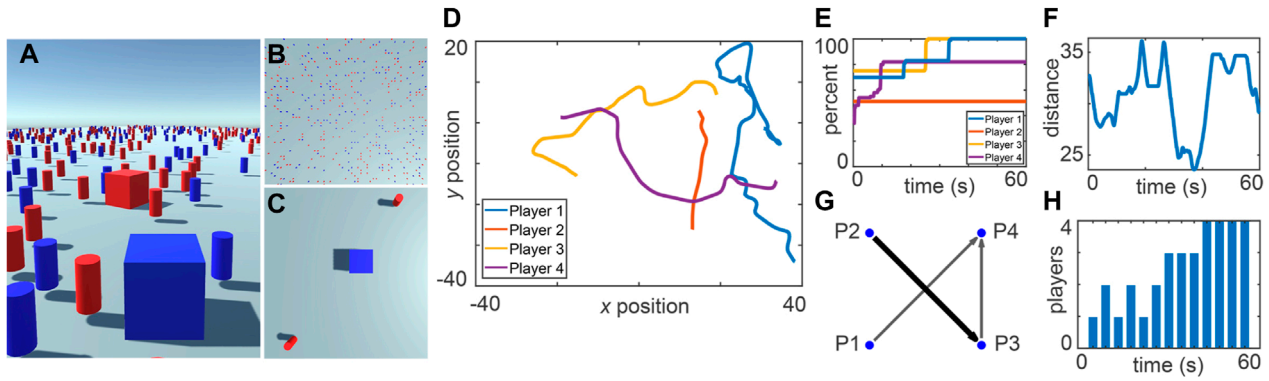


FIGURE 1
Collective decision making. Participants control cubic avatars while having either a broad **(A)** or limited **(C)** view of the full environment **(B)**. A wide variety of variables can be measured during the experiment, such as the participants' trajectories **(D)**, the percentage of the environment they have explored **(E)**, the average distance between participants **(F)**, the participant-participant viewing network **(G)**, and the number of avatars choosing the correct color **(H)**. Figure reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer eBook, Coucke et al. (2020), © Springer Nature 2020.

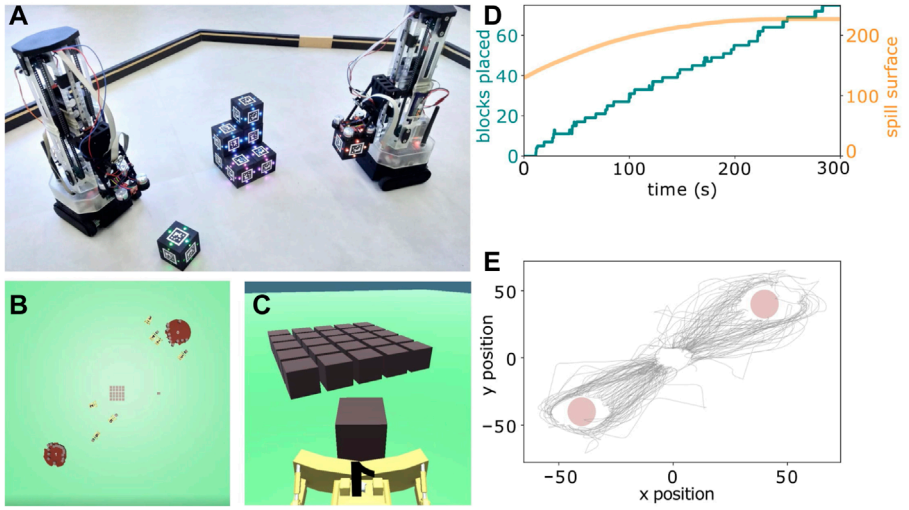


FIGURE 2
Collective construction. **(A)** Two physical robots that perform collective construction using stigmergic blocks (Allwright et al., 2019). Figure **(A)** reprinted from Allwright et al. (2019) under license CC BY-NC-ND 4.0. **(B, C)** "Lava spill task" in which participants use indirect communication to collectively construct a barrier to contain expanding spills. **(D)** The spill size stagnates after around 200 s, when participants successfully enclosed it with construction blocks. **(E)** Data such as the avatar trajectories can be used to analyze how participants coordinate the placement of blocks. Figures **(B–D)** adapted from Coucke et al. (2021) under license CC BY 4.0.

embodied scenarios that would be relevant to robots, in which, e.g., exploration and signalling can take place simultaneously. In our example implementation in Coucke et al. (2020), each of four participants controls the movements of a cubic avatar in an environment scattered with red and blue cylindrical landmarks (see Figure 1). The task is to explore the environment while making observations through the avatar's (broad or limited) field of view and simultaneously deciding whether there are more red or blue landmarks present in the environment. The participants must come to a consensus in order to complete the task and are only permitted to communicate with each other indirectly: they vote by changing their avatar color and they observe the avatar colors of the other participants while making their decisions (see Figures 1A–C). During an experiment, all perceptual information available to each participant, along with their actions, are recorded at a sampling rate of 10 Hz (Figures 1D–H).

In this experiment setup, participants came to a consensus about the predominant color in the environment through a combination of environmental and social information. In the example trial shown in Figure 1, at 45 s, all four participants had adopted the correct opinion (Figure 1H) after individually and broadly exploring the environment and then reducing their average relative distances to increase their access to social information (Figure 1F) and finally come to a consensus. When a consensus was reached, not all participants had personally observed all parts of the environment (Figure 1E), implying that social information was effectively used. Further, all participants had a strong directional line-of-sight connection with

at least one other participant (**Figure 1G**), but the most looked-at participant (P4) had not personally observed the whole environment (**Figure 1E**), implying that the consensus on the correct opinion was indeed arrived at by a self-organized and collective process. For more information on this and similar experiments, please refer to Coucke et al. (2020). By setting up more advanced experiments in this direction, data could be collected to, for example, identify social heuristics that can inspire new protocols in future robot swarms.

Collective construction

Existing swarm robotics approaches to construction often use stigmergy (i.e., indirect communication through modification of the environment) to coordinate (Petersen et al., 2019), but the structures built strictly by stigmergy are relatively simple. Future robot swarms should be able to build complex structures in dynamically changing environments (Dorigo et al., 2020). In our example "lava spill task" scenario in Coucke et al. (2021), human social behaviors in collective construction scenarios can be observed. In this task (see **Figure 2**), participants are instructed to collectively construct a barrier to contain an expanding spill, but are not instructed how to coordinate. Each participant controls the movement of an avatar that can push construction blocks. The environment includes two different spills (i.e., expanding circles) and a supply of construction blocks placed in between them. During an experiment, a group of eight participants needs to assess the environment and coordinate their actions using indirect communication (i.e., observation of peers) to barricade both of the expanding spills within 300 s.

The avatar trajectories in **Figure 2E** show that participants coordinated to distribute their work between the two spills and place blocks around the full circumferences of both spills. **Figure 2D** shows that participants continued to place more blocks at a roughly constant rate throughout the experiment, implying that no bottleneck arose in their self-organized coordination. The figure also shows that the expansion of both spills had successfully been stopped at around 200 s. For more information on this and similar experiments, please refer to Coucke et al. (2021). Using more advanced setups in this direction, the gathered behavioral data could provide insights into how self-organized coordination and group actions unfold over time and adapt to the environment. In order to get detailed information about participant strategies, experiments in this virtual environment can be temporarily interrupted at certain times to ask participants about, e.g., their explicit judgements about the beliefs of other participants, their sense of (joint) agency, or their feelings of alignment with others.

Discussion

Some features of human social groups, such as collective intentions, reflective discussion, or shared biases, might at first seem not particularly relevant for robots. However, there are many autonomous group behaviors that have not yet been demonstrated in self-organized robots. For instance, it is not yet understood how to have robots autonomously identify when they should make a collective decision (Khaluf et al., 2019). These fundamentals of group-level autonomy, which social animals such as humans exhibit effortlessly and consistently, might possibly be based on, or even depend on, such unexpected features as shared biases. Our perspective is that research that investigates the transfer of such social traits from humans to robots can help us to identify and understand the basic elements needed to build artificial social cognition.

Artificial restrictions in embodied experiments are unlikely to reveal how humans would behave in natural conditions, but there is existing evidence that such restrictions indeed have the potential to reveal aspects of embodied human social behavior that would be transferable to robots. For example, when realistic social cues such as gaze and facial expressions are inhibited, humans have been shown to focus on other communication channels, such as implicit movement-based communication (Roth et al., 2016).

If eventually achieved, the creation of social cognition among robots would open many further research questions. For instance, there are human collective intentions that go beyond the humans that are immediately present (Tomasello et al., 2005)—if robots have advanced social cognition abilities, how should different social groups of robots interact with each other, whether physically or remotely? As another example, intrinsic motivation or curiosity-driven learning could be investigated to motivate agents to explore the complex internal states that make up another agent, perhaps constituting a rudimentary theory of an artificial mind. Or, perhaps robots could be intrinsically motivated to autonomously develop completely new forms of artificial social cognition that do not resemble those already seen in humans or social animals.

Data availability statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the ethical committee of the Université libre de Bruxelles (permission 126/2020). The patients/participants provided their written informed consent to participate in this study.

Author contributions

All the authors contributed the ideas and concepts presented in the paper. NC and MKH wrote the first draft of the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

Funding

This work was supported by the program of Concerted Research Actions (ARC) of the Université libre de Bruxelles.

Acknowledgments

MKH, AC, and MD acknowledge support from the F.R.S.-FNRS, of which they are, respectively, postdoctoral researcher and research directors.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Albrecht, S. V., and Stone, P. (2018). Autonomous agents modelling other agents: A comprehensive survey and open problems. *Artif. Intell.* 258, 66–95. doi:10.1016/j.artint.2018.01.002

Allwright, M., Zhu, W., and Dorigo, M. (2019). An open-source multi-robot construction system. *HardwareX* 5, e00050. doi:10.1016/j.ohx.2018.e00050

Almaatouq, A., Noriega-Campero, A., Alotaibi, A., Krafft, P. M., Moussaid, M., and Pentland, A. (2020). Adaptive social networks promote the wisdom of crowds. *Proc. Natl. Acad. Sci.* 117, 11379–11386. doi:10.1073/pnas.1917687117

Baker, C. L., Saxe, R., and Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition* 113, 329–349. doi:10.1016/j.cognition.2009.07.005

Bard, N., Foerster, J. N., Chandar, S., Burch, N., Lanctot, M., Song, H. F., et al. (2020). The hanabi challenge: A new frontier for ai research. *Artif. Intell.* 280, 103216. doi:10.1016/j.artint.2019.103216

Bolotta, S., and Dumas, G. (2022). Social neuro AI: Social interaction as the "dark matter" of AI. *Front. Comput. Sci.* 4. doi:10.3389/fcomp.2022.846440

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi:10.1037/0033-295x.108.3.624

Buzsáki, G. (2019). *The brain from inside out*. Oxford University Press. doi:10.1093/oso/9780190905385.001.0001

Cangelosi, A., and Asada, M. (2022). *Cognitive robotics*. MIT Press.

Coucke, N., Heinrich, M. K., Cleeremans, A., and Dorigo, M. (2020). "Hugos: A multi-user virtual environment for studying human–human swarm intelligence," in *ANTS 2020–International conference on swarm intelligence* (Springer), 161–175. doi:10.1007/978-3-030-60376-2_13

Coucke, N., Heinrich, M. K., Cleeremans, A., and Dorigo, M. (2021). HuGoS: A virtual environment for studying collective human behavior from a swarm intelligence perspective. *Swarm Intell.* 15, 339–376. doi:10.1007/s11721-021-00199-1

Dale, R., Fusaroli, R., Duran, N. D., and Richardson, D. C. (2013). The self organization of human interaction. *Psychol. Learn. Motivation* 59, 43–95.

Daniel, C., Kroemer, O., Viering, M., Metz, J., and Peters, J. (2015). Active reward learning with a novel acquisition function. *Aut. Robots* 39, 389–405. doi:10.1007/s10514-015-9454-z

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Trans. R. Soc. B Biol. Sci.* 362, 679–704. doi:10.1098/rstb.2006.2004

De Jaegher, H., Di Paolo, E., and Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends cognitive Sci.* 14, 441–447. doi:10.1016/j.tics.2010.06.009

Dorigo, M., Theraulaz, G., and Trianni, V. (2020). Reflections on the future of swarm robotics. *Sci. Robotics* 5, eabe4385. doi:10.1126/scirobotics.abe4385

Dorigo, M., Theraulaz, G., and Trianni, V. (2021). Swarm robotics: Past, present, and future [point of view]. *Proc. IEEE* 109, 1152–1165. doi:10.1109/jproc.2021.3072740

Dumas, G., and Fairhurst, M. T. (2021). Reciprocity and alignment: Quantifying coupling in dynamic interactions. *R. Soc. Open Sci.* 8. doi:10.1098/rsos.210138

Friston, K., and Frith, C. (2015). A duet for one. *Conscious. Cognition* 36, 390–405. doi:10.1016/j.concog.2014.12.003

Frith, C. D. (2008). Social cognition. *Philosophical Trans. R. Soc. B Biol. Sci.* 363, 2033–2039. doi:10.1098/rstb.2008.0005

Frith, C. D., and Frith, U. (2012). Mechanisms of social cognition. *Annu. Rev. Psychol.* 63, 287–313. doi:10.1146/annurev-psych-120710-100449

Frohnwieser, A., Murray, J. C., Pike, T. W., and Wilkinson, A. (2016). Using robots to understand animal cognition. *J. Exp. analysis Behav.* 105, 14–22. doi:10.1002/jeab.193

Gordon, J., Maselli, A., Lancia, G. L., Thiery, T., Cisek, P., and Pezzulo, G. (2021). The road towards understanding embodied decisions. *Neurosci. Biobehav. Rev.* 131, 722–736. doi:10.1016/j.neubiorev.2021.09.034

Haggard, P., and Chambon, V. (2012). Sense of agency. *Curr. Biol.* 22, R390–R392. doi:10.1016/j.cub.2012.02.040

He, H., Boyd-Graber, J., Kwok, K., and Daumé, H.III (2016). "Opponent modeling in deep reinforcement learning," in *International conference on machine learning* (New York: PMLR), 1804–1813.

Heinrich, M. K., Wahby, M., Dorigo, M., and Hamann, H. (2022). "Swarm robotics," in *Cognitive robotics*. Editors A. Cangelosi, and M. Asada (MIT Press), 77–98.

Henschel, A., Hortensius, R., and Cross, E. S. (2020). Social cognition in the age of human–robot interaction. *Trends Neurosci.* 43, 373–384. doi:10.1016/j.tins.2020.03.013

Hertwig, R., and Herzog, S. M. (2009). Fast and frugal heuristics: Tools of social rationality. *Soc. Cogn.* 27, 661–698. doi:10.1521/soco.2009.27.5.661

Kameda, T., Toyokawa, W., and Tindale, R. S. (2022). Information aggregation and collective intelligence beyond the wisdom of crowds. *Nat. Rev. Psychol.* 1, 345–357. doi:10.1038/s44159-022-00054-y

Kengyel, D., Hamann, H., Zahadat, P., Radspieler, G., Wotawa, F., and Schmickl, T. (2015). "Potential of heterogeneity in collective behaviors: A case study on heterogeneous swarms," in *Prima 2015: Principles and practice of multi-agent systems*. Editors Q. Chen, P. Torroni, S. Villata, J. Hsu, and A. Omicini (Cham: Springer International Publishing), 201–217. doi:10.1007/978-3-319-25524-8_13

Khaluf, Y., Simoens, P., and Hamann, H. (2019). The neglected pieces of designing collective decision-making processes. *Front. Robotics AI* 6, 16. doi:10.3389/frobt.2019.00016

Lokesh, R., Sullivan, S., Calalo, J. A., Roth, A., Swanik, B., Carter, M. J., et al. (2022). Humans utilize sensory evidence of others' intended action to make online decisions. *Sci. Rep.* 12, 8806. doi:10.1038/s41598-022-12662-y

Mathews, N., Christensen, A. L., O'Grady, R., Mondada, F., and Dorigo, M. (2017). Mergeable nervous systems for robots. *Nat. Commun.* 8, 439–447. doi:10.1038/s41467-017-00109-2

Moussaid, M., Helbing, D., and Theraulaz, G. (2011). How simple rules determine pedestrian behavior and crowd disasters. *Proc. Natl. Acad. Sci.* 108, 6884–6888. doi:10.1073/pnas.1016507108

Murata, S., Yamashita, Y., Arie, H., Ogata, T., Sugano, S., and Tani, J. (2015). Learning to perceive the world as probabilistic or deterministic via interaction with others: A neuro-robotics experiment. *IEEE Trans. neural Netw. Learn. Syst.* 28, 830–848. doi:10.1109/tnnls.2015.2492140

Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818. doi:10.1038/nn1911

Nouyan, S., Groß, R., Bonani, M., Mondada, F., and Dorigo, M. (2009). Teamwork in self-organized robot colonies. *IEEE Trans. Evol. Comput.* 13, 695–711. doi:10.1109/TEVC.2008.2011746

Olsson, A., Knapska, E., and Lindström, B. (2020). The neural and computational systems of social learning. *Nat. Rev. Neurosci.* 21, 197–212. doi:10.1038/s41583-020-0276-4

Patel, D., Fleming, S. M., and Kilner, J. M. (2012). Inferring subjective states through the observation of actions. *Proc. R. Soc. B* 279, 4853–4860. doi:10.1098/rspb.2012.1847

Petersen, K. H., Napp, N., Stuart-Smith, R., Rus, D., and Kovac, M. (2019). A review of collective robotic construction. *Sci. Robotics* 4, eaau8479. doi:10.1126/scirobotics.aau8479

Pezzulo, G., and Dindo, H. (2011). What should i do next? Using shared representations to solve interaction problems. *Exp. Brain Res.* 211, 613–630. doi:10.1007/s00221-011-2712-1

Rabinowitz, N., Perbet, F., Song, F., Zhang, C., Eslami, S. M. A., and Botvinick, M. (2018). "Machine theory of mind," in Proceedings of the 35th International Conference on Machine Learning Editors J. Dy, and A. Krause (Stockholm: PMLR), 80, 4218–4227.

Raileanu, R., Denton, E., Szlam, A., and Fergus, R. (2018). "Modeling others using oneself in multi-agent reinforcement learning," in *Proceedings of the 35th international conference on machine learning* Editors J. Dy, and A. Krause (Stockhom: PMLR) 80, 4257–4266.

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi:10.1162/neco.2008.12-06-420

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi:10.1146/annurev.neuro.27.070203.144230

Roth, D., Lugrin, J.-L., Galakhov, D., Hofmann, A., Bente, G., Latoschik, M. E., et al. (2016). "Avatar realism and social interaction quality in virtual reality," in *2016 IEEE virtual reality (VR)* (IEEE), 277–278. doi:10.1109/vr.2016.7504761

Rubenstein, M., Cornejo, A., and Nagpal, R. (2014). Programmable self-assembly in a thousand-robot swarm. *Science* 345, 795–799. doi:10.1126/science.1254295

Saxe, R. (2005). Against simulation: The argument from error. *Trends Cognitive Sci.* 9, 174–179. doi:10.1016/j.tics.2005.01.012

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., et al. (2013). Toward a second-person neuroscience. *Behav. Brain Sci.* 36, 393–414. doi:10.1017/s0140525x12000660

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends Cognitive Sci.* 10, 70–76. doi:10.1016/j.tics.2005.12.009

Seymour, B., and Dolan, R. (2008). Emotion, decision making, and the amygdala. *Neuron* 58, 662–671. doi:10.1016/j.neuron.2008.05.020

Silver, C. A., Tatler, B. W., Chakravarthi, R., and Timmermans, B. (2020). Social agency as a continuum. *Psychonomic Bull. Rev.* 28, 434–453. doi:10.3758/s13423-020-01845-1

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691. doi:10.1017/s0140525x05000129

Tump, A. N., Pleskac, T. J., and Kurvers, R. H. J. M. (2020). Wise or mad crowds? The cognitive mechanisms underlying information cascades. *Sci. Adv.* 6, eabb0266. doi:10.1126/sciadv.abb0266

Ullman, T., Baker, C., Macindoe, O., Evans, O., Goodman, N., and Tenenbaum, J. (2009). "Help or hinder: Bayesian models of social goal inference," in *Advances in neural information processing systems*. Editors Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta (Vancouver: Curran Associates, Inc.), 22, 1874–1882.

Valentini, G., Ferrante, E., and Dorigo, M. (2017). The best-of-n problem in robot swarms: Formalization, state of the art, and novel perspectives. *Front. Robotics AI* 4. doi:10.3389/frobt.2017.00009

Valentini, G., Ferrante, E., Hamann, H., and Dorigo, M. (2016). Collective decision with 100 kilobots: Speed versus accuracy in binary discrimination problems. *Aut. agents multi-agent Syst.* 30, 553–580. doi:10.1007/s10458-015-9323-3

Van Vugt, M. (2006). Evolutionary origins of leadership and followership. *Personality Soc. Psychol. Rev.* 10, 354–371. doi:10.1207/s15327957pspr1004_5

Vernon, D. (2014). *Artificial cognitive systems: A primer*. MIT Press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bull. Rev.* 9, 625–636. doi:10.3758/bf03196322

Wykowska, A., Chaminade, T., and Cheng, G. (2016). Embodied artificial agents for understanding human social cognition. *Philosophical Trans. R. Soc. B Biol. Sci.* 371, 20150375. doi:10.1098/rstb.2015.0375

Zhu, W., Allwright, M., Heinrich, M. K., Oğuz, S., Christensen, A. L., and Dorigo, M. (2020). "Formation control of uavs and mobile robots using self-organized communication topologies," in *ANTS 2020–International conference on swarm intelligence* (Springer), 306–314. doi:10.1007/978-3-030-60376-2_25

frontiers | Frontiers in Neurorobotics

TYPE Original Research
PUBLISHED 20 June 2023
DOI 10.3389/fnbot.2023.1181143

OPEN ACCESS

EDITED BY
Georg Northoff,
University of Ottawa, Canada

REVIEWED BY
Wenli Zhang,
Beijing University of Technology, China
Stefania Costantini,
University of L'Aquila, Italy

*CORRESPONDENCE
Fanliang Bu
✉ bufanliang@sina.com

RECEIVED 07 March 2023
ACCEPTED 26 May 2023
PUBLISHED 20 June 2023

CITATION
Zhai H, Lv X, Hou Z, Tong X and Bu F (2023) MLNet: a multi-level multimodal named entity recognition architecture.
Front. Neurorobot. 17:1181143.
doi: 10.3389/fnbot.2023.1181143

MLNet: a multi-level multimodal named entity recognition architecture

Hanming Zhai[1], Xiaojun Lv[2], Zhiwen Hou[1], Xin Tong[1] and Fanliang Bu[1]*

[1]School of Information Network Security, People's Public Security University of China, Beijing, China, [2]Institute of Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing, China

In the field of human–computer interaction, accurate identification of talking objects can help robots to accomplish subsequent tasks such as decision-making or recommendation; therefore, object determination is of great interest as a pre-requisite task. Whether it is named entity recognition (NER) in natural language processing (NLP) work or object detection (OD) task in the computer vision (CV) field, the essence is to achieve object recognition. Currently, multimodal approaches are widely used in basic image recognition and natural language processing tasks. This multimodal architecture can perform entity recognition tasks more accurately, but when faced with short texts and images containing more noise, we find that there is still room for optimization in the image-text-based multimodal named entity recognition (MNER) architecture. In this study, we propose a new multi-level multimodal named entity recognition architecture, which is a network capable of extracting useful visual information for boosting semantic understanding and subsequently improving entity identification efficacy. Specifically, we first performed image and text encoding separately and then built a symmetric neural network architecture based on Transformer for multimodal feature fusion. We utilized a gating mechanism to filter visual information that is significantly related to the textual content, in order to enhance text understanding and achieve semantic disambiguation. Furthermore, we incorporated character-level vector encoding to reduce text noise. Finally, we employed Conditional Random Fields for label classification task. Experiments on the Twitter dataset show that our model works to increase the accuracy of the MNER task.

KEYWORDS

multimodal named entity recognition, short text, multi-head attention, pre-training, cross task

1. Introduction

In the field of human–computer interaction, a large number of novel techniques have been proposed to improve efficiency, reduce operational difficulty, and increase recognition accuracy. Currently, natural language-based human–robot interaction has been widely used (Ahn et al., 2018; Park et al., 2019; Walker et al., 2019). At the same time, the field of image perception is also developing rapidly. The combination of visual information and natural language can further improve the service response capability of robots. Especially when robots need to perform tasks related to entity features, it is meaningful to introduce multimodal named entity recognition techniques. The human-generated text that the robot needs to process is often spoken and noisy, which is somewhat similar to the characteristics of free posting on social media. Therefore, we choose the corpus and images from social

media postings for the named entity recognition task, thus helping to improve multimodal human–robot interaction.

With the widespread use of social media platforms, the number of individual user postings has grown rapidly. Such interesting and diverse informal expressions provide users with rich information while providing a large amount of raw corpus data for natural language processing (NLP). Named Entity Recognition (NER), as a precursor to many information extraction tasks, aims to discover multiple categories of named entities, such as Person (PER), Location (LOC), and Organization (ORG), from raw text data. Given its importance, NER has attracted significant attention in the research community (Lample et al., 2016; Jiang et al., 2021; Radmard et al., 2021; Tian et al., 2021).

Although a large number of excellent methods have emerged that enable increasing accuracy in named entity recognition efforts, most of them are based on news report texts or domain-length texts (Li et al., 2021; Wang et al., 2021, 2022). When solving named entity recognition tasks for social media texts (e.g., tweets), their shorter length and large amount of noise are fully considered, and thus, their performance is often much lower than that in news report texts. In general, common text noise in tweets includes misspellings, web abbreviations, and some informal newly invented words. In recent years, the posting format of social media platforms has been innovated, and the "text-image" format has gradually become mainstream, and some studies have proposed using visual features to improve the performance of NER (Arshad et al., 2019; Asgari-Chenaghlu et al., 2020; Chen et al., 2022). In this study, we will focus on multimodal named entity recognition (MNER) for social media postings, with the goal of extracting the corresponding entities from "text-image" pairs. This is shown in Figure 1.

A large number of excellent algorithms for multimodal named entity recognition already exist (Moon et al., 2018; Chen D. et al., 2021), but there are still some problems as follows: (1) The expressions in social media posts are often informal, colloquial, and even have certain spelling errors, which can affect the accuracy of text recognition. (2) Since short texts in social media contain less contextual information, there is some difficulty in determining entity types, which requires the use of image information to achieve semantic disambiguation. (3) For image information screening and fusion problems, there may be a large amount of irrelevant information in the whole picture, and there is some interference in entity extraction. Even the accompanying images in some posts may be irrelevant to the text, resulting in lower accuracy when directly fusing image features with text.

To solve these problems, we will pre-train the image and text data separately to extract the focused object features in the image and the word embedding and character embedding in the text as the input. We propose a multi-modal BERT model that uses a filtering gate to preserve visually relevant image features and trains an adaptive attention network to fuse image and textual features. Specifically, we design a symmetrical image-text fusion module that fully integrates multi-modal data by exchanging filtered feature information bi-directionally. To address the limited expressiveness and noise issues of short textual information, we further enhance the understanding of textual features using a bidirectional long short-term memory network. Finally, we use a conditional random field for label classification and complete the entity recognition task. We conduct experiments on the Twitter dataset to validate the effectiveness of our proposed model.

The main contributions of this study are as follows.

1. Based on a more efficient multimodal encoding scheme to reduce the noise in social media posts and improve the accuracy of named entity recognition tasks in informal corpus;
2. Proposes a new symmetric unified architecture for named entity recognition based on image-text features for multi-channel and multi-level input computation;
3. Achieves better performance on the Twitter2015 and Twitter2017 datasets.

2. Related work

2.1. Named entity recognition

NER has been attracting a lot of attention from the research community as a precursor to many natural language processing (NLP) tasks. While traditional NER tasks often rely on specific knowledge and manual annotation combined with statistical

A beautiful night time photo of the [Sydney Opera house **MISC**] in [Sydney Australia **LOC**]. (Photo by [Glen Anderson **PER**])

[Kenya Airways **ORG**] posts record Sh25.7bn net loss, compared to loss of Sh3.3bn in previous financial year

FIGURE 1
Two examples for multimodal named entity recognition.

learning methods, with the continuous development of neural networks, most NERs are now opting for deep learning approaches. Various supervised learning-based approaches have been proposed, focusing primarily on designing the structure of neural networks so that more valuable features are fed into the classifier. The first use of neural networks applied to the study of named entities was by Hammerton (2003). They used a one-way long short-term memory network (LSTM), which has good sequence modeling capabilities, and LSTM-CRF became the underlying architecture for entity recognition. Later, based on this model, Lample et al. (2016) proposed a neural network model combining Bidirectional Long Short-Term Memory (BiLSTM) and Conditional Random Fields (CRF), a bidirectional structure capable of acquiring contextual sequence information. Pinheiro and Collobert (2014) first applied the combination of CNN and CRF structures in named entity recognition research with good results on CoNLL-2003. In the field of chemistry, Luo et al. (2018) used the BiLSTM-CRF model based on the attention mechanism to further improve the performance of entity recognition.With the introduction of the BERT model, more research has focused on the improvement and optimization of the pre-trained model, and the effectiveness of the NER task has been further improved (Devlin et al., 2018; Jawahar et al., 2019; Souza et al., 2019). The combination of BERT as an encoder with LSTM-CRF by Liu et al. (2019) and Luo et al. (2020) and the combination of BERT and a threshold control unit (GRU) by Liu et al. resulted in better results for the NER task on the CoNLL-2003 dataset.

However, these methods tend to be more suitable for NER tasks in formal texts, and most of them do not achieve satisfactory results when facing social media texts. To address this problem, many studies have added some resources other than text to the input of NER, which can achieve better performance on social media texts. Moreover, as the image-text posting format on social media becomes mainstream (Su et al., 2019; Tan and Bansal, 2019), recent studies tend to focus more on multimodal named entity recognition tasks (MNER), to improve the accuracy of text feature extraction with the help of feature elements in images.

2.2. Multimodal named entity recognition

Moon et al. (2018) transformed the NER task into a sequence annotation problem with character embeddings and word embeddings as text data inputs and a weighted combination of textual and visual information through an adaptive co-attentive network. Lu et al. (2018) proposed a salient visual attention model to find image regions associated with textual content. The visual contextual information is extracted and fused into the word-level output of the biLSTM model. Zhang et al. (2018) designed an adaptive co-attentive network layer to simultaneously learn the fused feature vectors of vision and language. Moreover, a filtering gate was designed to determine whether the fused features contribute to the accuracy of the annotation task. Arshad et al. (2019) also proposed a gated multimodal fusion representation, where gated fusion is a weighted sum of visual attention features and marker alignment features. Whereas, additive attention scores between word queries and picture features were used to weigh and calculate the visual attention features using VGG-19 visual features (Simonyan and Zisserman, 2014). Chen S. et al. (2020) and Chen X. et al. (2021) extracted visual information into subtitles and proposed a softer method of image-text combination that improves the fusion of different modal features.

2.3. Visual-linguistic pre-training

Visual-linguistic pre-training mainly learns the semantic correspondence between different modalities by pre-training on large-scale data to achieve proper operation of the model in resource-poor scenarios. There are already many models implemented based on single-stream (Murahari et al., 2020; Hong et al., 2021) and dual-stream (Gan et al., 2020; Gao et al., 2020) architectures and pre-trained on common datasets such as COCO (Lin et al., 2014). Many VLP models use a single encoder architecture, where the multimodal fusion representation is fed directly to the output layer to generate the final output. In contrast, other VLP models advocate the use of the encoder–decoder architecture commonly used by Transformer, where the multimodal representation is first fed to the decoder and then to the output layer.

Overall, the key issues of the current multimodal named entity recognition task remain in the strategy of combining visual features with linguistic features (Li et al., 2022) and the subsequent entity extraction architecture (He et al., 2016; Li et al., 2020). Therefore, in this study, we explore a multi-channel and multi-level named entity recognition architecture to accomplish MNER for text-image modal tweets.

3. Methodology

The overall structure of the MLNet model roughly consists of several components (shown in Figure 2): image feature extraction, text embedding, multimodal fusion embedding, and label prediction. We first perform image feature extraction, and subsequently input the overall image features and focused relevant visual information into the multimodal BERT along with the text vector to obtain a sequence vector containing rich entity information. Subsequently, contextual information is further learned in the BiLSTM structure. CRF are used to perform the final label prediction.

3.1. Image feature extraction

In the case where an image is associated with the corresponding text content, the named entities mentioned in the text are often associated with only the salient features in the image. For the original images in the dataset, the top s salient local visual objects in them are first extracted using a pre-trained target detection model. Then, we align the global and local images by scaling so that the size of each image is 224 × 224 pixels, noted as the global image L_0 and local image $L = \{L_1, L_2, ..., L_s\}$.

Features were subsequently extracted from the images L and L_0 using ResNet as the image feature input for the entire subsequent

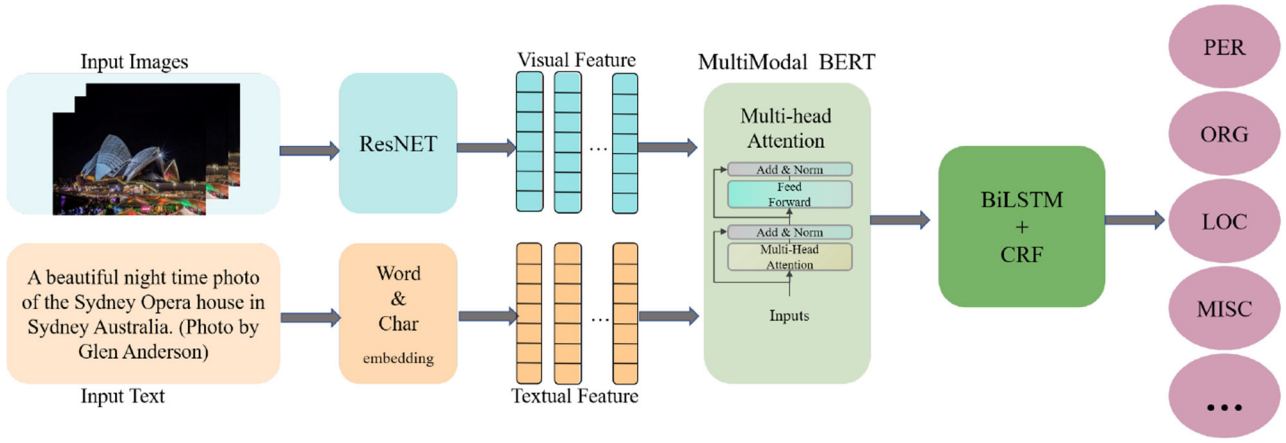


FIGURE 2
Overall architecture of our MLNet.

named entity recognition network architecture. Considering the excellent performance of the ResNet network in object detection tasks, we hypothesize that residual networks can also effectively recognize entities that may appear in images, thereby reducing the impact of image noise on entity recognition accuracy. We use a lightweight ResNet structure with 18 convolutional layers, where each layer uses a convolutional kernel, and residual blocks are used between each convolutional layer to increase the depth of the model. The last pooling layer feature in ResNet is retained with dimension $7 \times 7 \times d$, where 7×7 is the 49 regions of the image and $d = 2048$ is the dimension of each visual region. At this point, each image can be represented as $\widetilde{v}_I = \{\widetilde{v}_i | \widetilde{v}_i \in \mathbb{R}, i = 1, 2, ..., 49\}$, and it is subsequently linearly varied and normalized to maintain the same dimensionality as the text vector.

$$v_I = tanh(W_I \cdot \widetilde{v}_I + b_I) \tag{1}$$

where the parameters W_I and b_I can be obtained from the training data. The image features obtained at this point can be used as input to the subsequent overall recognition framework to be combined with text features. The visual feature pre-processing process is shown in Figure 3.

3.2. Image-text feature fusion

Taking into account both the complexity of the model and the actual effectiveness, we chose the WordPiece tokenizer used in BERT for tokenization and word embedding. For character embedding, we adopted a basic recurrent neural network (RNN) model. After completing image and text data pre-processing separately, we choose a BERT structure capable of handling multimodal inputs to realize the joint representation of text and images. The text vectors in the dataset (which have completed word embedding and character embedding) and their corresponding set of image vectors (including global images and local images) are sent to the image converter and text converter independently as inputs to the multimodal BERT. The two converters are trained separately and do not share parameters, and the converters are implemented by a multi-head attention mechanism as follows:

$$MultiHead(Q_F, K_F, V_F) = Concat(head_1^F, ..., head_n^F)W^O \tag{2}$$

$$Head_i^F = Attention(QW_i^{Q_F}, KW_i^{K_F}, VW_i^{V_F}) \tag{3}$$

where Q is the query matrix, K is the key matrix, V is the value matrix, W is the weight matrix, and $F = \{Text, Image\}$ is used to distinguish the image conversion module from the text conversion module. The multi-head attention mechanism projects Q, K, and V through several different linear transformations, and finally, stitches the different attention results together. To improve the task efficiency, we add a cross-attention layer for cross-modal interaction and introduce a control unit to discard the visual information that is less relevant to the text. The main structure of our multimodal BERT is shown in Figure 4.

For the multimodal BERT structure, we choose three common target tasks for pre-training as follows: (i) masked language modeling (MLM); (ii) masked vision modeling (MVM); and (iii) visual-language matching (VLM). In task, (i) some elements of the text input are masked, but vectors corresponding to image regions are not masked. Task (ii), in contrast to (i), partially masks the image input, but the corresponding text vectors are not masked. In the MVM task, we mark all masked pixel values in the image as a special value of -1 to distinguish them from the original colored image.

The MLM task can be expressed as follows:

$$\mathfrak{L}_{MLM} = -E_{(\mathbf{v},\mathbf{w})\sim D} logP(\mathbf{w}_m | \mathbf{w}_{\not m}, \mathbf{v}) \tag{4}$$

where $\mathbf{v}$ denotes visual, $\mathbf{w}$ denotes text, $\mathbf{w}_m$ denotes masked text, $\mathbf{w}_{\not m}$ denotes remained text, and D denotes the training data set.

The MVM task can be expressed as follows:

$$\mathfrak{L}_{MVM} = E_{(\mathbf{v},\mathbf{w})\sim D} f(\mathbf{v}_m | \mathbf{v}_{\not m}, \mathbf{w}) \tag{5}$$

We choose label classification as the MVM pre-training task, where the mask features are fed into the FC layer to predict

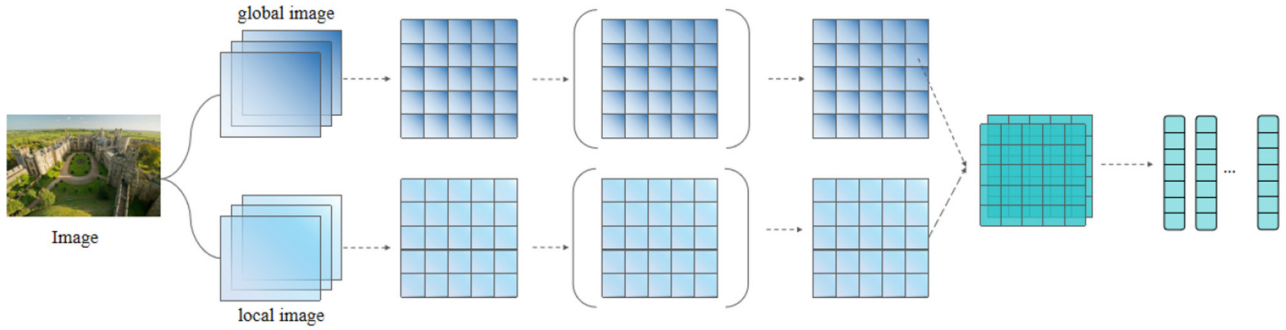


FIGURE 3
Image feature extraction.

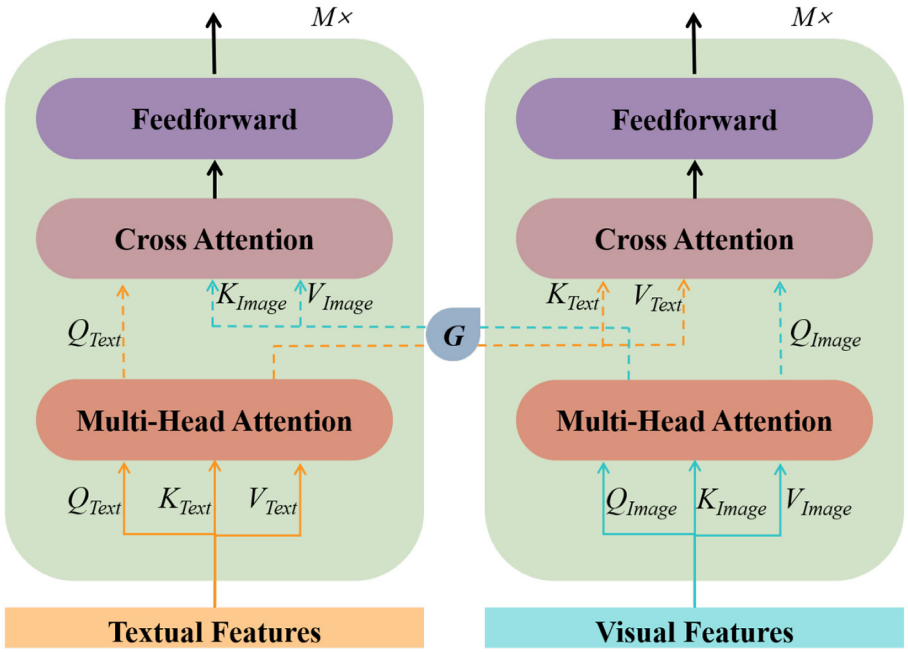


FIGURE 4
Multi-modal BERT architecture.

the category scores of the objects and then normalized using the softmax function. The detected object categories are used as labels for the masked regions as follows:

$$f_1(\mathbf{v}_m|\mathbf{v}_{\not m}, \mathbf{w}) = \sum_{i=1}^{K} CE(c(\mathbf{v}_m^i) - g_1(\mathbf{v}_m^i)) \tag{6}$$

where $g_1(\mathbf{v}_m^i)$ is the class of detected objects, and K denotes the number of visual areas.

We will minimize the KL dispersion (Chen Y.-C. et al., 2020) between the two distributions of the real object class and the detected object class to complete the supervised learning, which is the original output of the detector as follows:

$$f_2(\mathbf{v}_m|\mathbf{v}_{\not m}, \mathbf{w}) = \sum_{i=1}^{K} K_{DL}(\hat{c}(\mathbf{v}_m^i) - g_2(\mathbf{v}_m^i)) \tag{7}$$

where $g_2(\mathbf{v}_m^i)$ is the distribution of detected object classes.

By performing the above two tasks, we complete the training of the transformer for each of the image module and text modules. Before the cross-attention layer, we fuse the keys K and values V of the image and text separately to achieve more fine-grained feature interactions. We use fully connected layers to perform the fusion process. The fused representation of the two patterns is then provided to the FC layer and the sigmoid function to predict a score between 0 and 1, where 0 means visual and verbal mismatch and 1 means visual and verbal match. When the predicted value is less than 0.5, we discard the visual vector because it indicates that the mapping is not relevant enough to the text, and adding visual features at this point will reduce the accuracy of entity recognition.

$$G(K, T) = \begin{cases} 1\,, sigmoid(LN(Text, Image)) > 0.5 \\ 0\,, sigmoid(LN(Text, Image)) \leqslant 0.5 \end{cases} \tag{8}$$

where $LN(\cdot)$ denotes the fusion score of image and text, and G is the cross-attention layer coefficient.

3.3. Label prediction

Since the short text can contain less contextual information, we further use the BiLSTM structure for contextual encoding, in order to preserve the complete semantics as much as possible. The core of the LSTM consists of the following structures: the forgetting gate, the input gate, the output gate, and the memory cell; the common function of the input gate and the forgetting gate is to discard the useless information and pass the useful information to the next moment. The output of the whole structure is obtained by multiplying the output of the memory cell and the output of the output gate. Since the one-way LSTM model cannot handle the contextual information at the same time, the BiLSTM (Bidirectional Long-Short Term Memory) proposed by Graves A et al. The basic idea is to take forward and backward LSTM for each input sequence, respectively, and then, the outputs of the same moment are merged. Thus, for each moment, there corresponds to forward and backward information, which can be expressed as $h_t = \left[\overrightarrow{h_t}, \overleftarrow{h_t}\right]$, where $\overrightarrow{h_t}$ and $\overleftarrow{h_t}$ denote the forward and backward outputs of the bi-directional LSTM, respectively. In the named entity recognition task, BiLSTM is good at handling long-range textual information but cannot handle the dependencies between neighboring labels. Moreover, CRF can obtain an optimal prediction sequence by the relationship of neighboring labels, which can compensate for the shortcomings of BiLSTM.

For the input sequence $X = \{x_1, x_2, ..., x_n\}$, it is assumed that $P \in \mathbb{R}^{n\times k}$ is the output score matrix of the BiLSTM, where n is the input vector dimension, k is the number of labels, and the score of the $j-$th label of the $i-$th word. For the prediction sequence $Y = \{y_1, y_2, ..., y_n\}$, the formula to calculate its score is as follows:

$$s(X, Y) = \sum_{i=0}^{n} A_{y_i, y_{i+1}} + \sum_{i=1}^{n} P_{i, y_i} \tag{9}$$

Where A denotes the matrix of transferred scores, A_{ij} representing the scores transferred from label i to label j. The size of A is $k + 2$. The probability of generating the predicted sequence Y is as follows:

$$p(Y|X) = \frac{e^{s(X,Y)}}{\sum_{\widetilde{Y}\in Y_X} s\left(X, \widetilde{Y}\right)} \tag{10}$$

The likelihood function of the predicted sequence is obtained by taking the logarithm on both sides as follows:

$$ln\left(p\left(Y|X\right)\right) = s\left(X, Y\right) - ln\left(\sum_{\widetilde{Y}\in Y_X} s\left(X, \widetilde{Y}\right)\right) \tag{11}$$

where X denotes the true labeled sequence and Y denotes all possible labeled sequences. The output sequence of the maximum score is obtained after decoding as follows:

$$Y^{\circ} = \underset{\widetilde{Y}\in Y_X}{\operatorname{argmax}}\, s\left(X, \widetilde{Y}\right) \tag{12}$$

TABLE 1 The basic statistics of our two Twitter datasets.

Entity type	Twitter2015			Twitter2017		
	Train	Dev	Test	Train	Dev	Test
Person	2,217	552	1,816	2,943	626	621
Location	2,091	522	1,697	731	173	178
Organization	928	247	839	1,647	375	395
Miscellaneous	940	225	726	701	150	157
Total	6,176	1,546	5,078	6,049	1,324	1,351
Num of Tweets	4,000	1,000	3,257	3,373	723	723

4. Experiment

4.1. Experiment settings

This study uses two publicly available Twitter datasets, Twitter2015 and Twitter2017, constructed by Zhang et al. (2018) and Lu et al. (2018), respectively. These two datasets mainly include multimodal user posts posted on Twitter during 2014–2015 and 2016–2017. Table 1 shows the number of entities and multimodal tweet counts for each type in the training, development, and testing sets for both datasets.

Evaluation system: The common labeling systems for named entity identification are the BIO system, BIOE system, and BIOES system, and the BIO system is chosen in this study. The system has nine labels, namely, "O", "B-PER", "I-PER", "B-ORG", "I-ORG", "B-LOC", "I-LOC", "B-MISC", and "I-MISC".

In this study, recall R, precision P, and F1 values are used to judge the performance of the model, and each evaluation index is calculated as follows:

$$P = \frac{a}{B} \times 100\% \tag{13}$$

$$R = \frac{a}{A} \times 100\% \tag{14}$$

$$F_1 = \frac{2PR}{P + R} \times 100\% \tag{15}$$

Where a is the number of correctly identified entities, A is the total number of entities, and B is the number of identified entities.

4.2. Baseline

To exemplify the effectiveness of our new models, we selected several benchmark models for comparison. We first consider a representative set of text-based models as follows: 1) CNN-BiLSTM-CRF, a widely adopted NER neural network model that is an improvement in BiLSTM-CRF that replaces word embeddings with character-level word embeddings and CNN-based concatenation of character-level word representations for each word; 2) BERT- CRF which is a pioneering study that eliminates the heavy reliance on hand-crafted features and simply employs a bi-directional LSTM model and then uses CRF layers for the final prediction of each word.

TABLE 2 Performance comparison on our two TWITTER datasets.

Methods	Twitter2015			Twitter2017		
	Precision	Recall	F1	Precision	Recall	F1
CNN-BiLSTM-CRF	66.24	68.09	67.15	80.00	78.76	79.37
BERT-CRF	69.22	74.59	71.81	83.32	83.57	83.44
AdapCAN-BERT-CRF	69.87	74.59	72.15	85.13	83.20	84.10
VisualBERT	68.84	71.39	70.09	84.06	85.39	84.72
OCSGA	74.71	71.21	72.92	-	-	-
UMT	71.67	75.23	73.41	85.28	85.34	85.31
UMGF	74.49	75.21	74.85	86.54	84.50	85.51
HVPNet	73.87	76.82	75.32	85.84	**87.93**	86.87
MLNet (ours)	**75.73**	**76.85**	**76.28**	**87.36**	86.97	**87.17**

The best results in the table are highlighted in bold.

In addition, we compare other multimodal approaches for named entity recognition as follows: (1) AdapCAN-Bert-CRF (Zhang et al., 2018) which designs an adaptive co-attention network to induce visual representations of each word; (2) VisualBERT (Li et al., 2019) which differs from the above mentioned SOTA based mainly on co-attention approach, VisualBERT is a single-stream structure, which is a strong baseline for comparison; (3) OCSGA (Wu et al., 2020), a model that combines dense co-attentive networks (self-attentive and guided attention), to model the association between visual objects and textual entities and the intrinsic connections between objects or entities; (4) UMT (Yu et al., 2020) by adding a multi modality converter to achieve a unified architecture by adding an auxiliary entity span detection module; (5) UMGF (Zhang et al., 2021), using a multimodal graph fusion approach, captures various semantic relationships between multimodal semantic units (words and visual objects); (6) HVPNet (Chen et al., 2022), using dynamic threshold aggregation strategy to achieve hierarchical multiscale visual features as fused visual prefixes.

In the multimodal BERT structure, we choose 12 layers, each hidden layer has a size of 768 and 12 self-attentive heads. In the training process of the BiLSTM-CRF structure, the Adam optimizer is used, and the learning rate is chosen as 0.001. In addition, the LSTM dimension is set to 200, the batch size is 64, and the maximum vector length of text input is 128, to prevent an overfitting problem. Dropout is used in the input and output of BiLSTM, and the value is 0.5 (Bouthillier et al., 2015).

4.3. Main results

In Table 2, we report the precision (P), recall (R), and F1 score (F1) achieved by each method on the two Twitter datasets.

First, compared with the two text-based methods CNN-BiLSTM-CRF and BERT-CRF, it is clearly observed that our model outperforms the other methods on both datasets. It is clear that the inclusion of visual features does guide the NER model well in discovering named entities when solving the named entity recognition task for tweet text. Although the accompanying images of some tweets may not be directly related to the text content, to a certain extent, extracting the corresponding image information (shown as Figure 5) can address the ambiguity issues and irregular representations present in the text.

Second, our model outperforms OCSGA, UMT, and UMGF compared with existing multimodal methods, thus it can be shown that the visual information preprocessed by focused feature extraction is more helpful for text entity extraction task effect enhancement than the complete image. By drawing on the target detection algorithm in the image preprocessing process, we are able to focus more actively on the entity information originally present in the image in the subsequent label classification task, thus avoiding a large amount of irrelevant information interference. Moreover, compared with the pre-trained model VisualBERT in which images and complete text are directly input to BERT for encoding and decoding, our model also introduces character-level embedding to address spelling errors and noise of informal expressions and uses a multi-head attention mechanism to learn different levels and modalities of input information.

Finally, comparing all named entity recognition methods, we can see that our multimodal entity recognition architecture achieves the best results, 0.29 and 1.86% higher than the second best method, respectively, and outperforms the twitter2015 dataset in twitter2017. This indicates that our model is accurate and effective in small-sample scenarios. In this regard, we analyze that our model is more enough to better learn the association between visual features and text features in small sample scenarios, thus improving the accuracy of tag classification.

4.4. Further analysis

4.4.1. Ablation study

To understand the role of network structure in the model, we performed more experiments for some variants of our model. For each dataset, we compared the full model MLNet and two variants of the model MLNet w/o Early Fusion and MLNet w/o Mtla. The obtained knots are shown in Table 3.

MLNet: Complete multi-level multimodal named entity recognition model, including image-text pre-processing, multimodal feature fusion already and feature annotation structure.

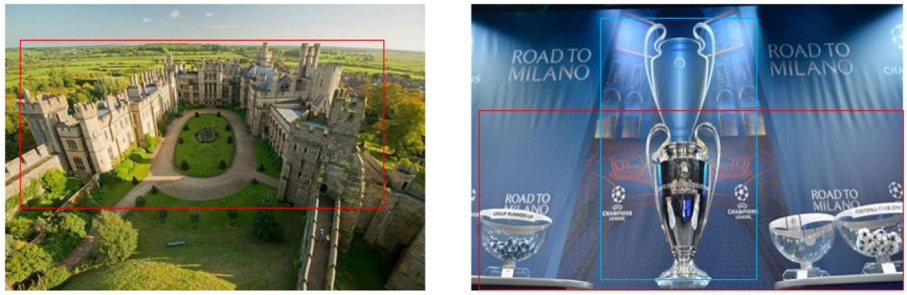


FIGURE 5
Two examples of correct visual attention. Our model successfully highlights related image regions required to predict correct tag.

TABLE 3 Ablation study of MLNet.

Methods	Twitter2015			Twitter2017		
	Precision	Recall	F1	Precision	Recall	F1
MLNet	**75.73**	**76.85**	**76.28**	**87.36**	**86.97**	**87.17**
MLNet w/o Fus.	74.44	73.07	73.75	85.10	84.51	84.80
MLNet w/o Mtla.	74.73	73.72	74.22	85.97	85.62	85.79

The best results in the table are highlighted in bold.

MLNet w/o Fus.: Instead of combining the image feature vectors extracted by ResNet with the text in the encoder Transformer in BERT, they are combined with the text vectors in the last layer of the Transformer structure before output. This allows us to test whether the interaction between linguistic and visual in the whole Transformer stack is important for performance. This variant structure is able to represent the impact of image features fused with text features for entity recognition tasks.

MLNet w/o Mtla.: When encoding text, only word-level vectors are kept, and no character-level feature vectors are introduced. This variant can help us to recognize the help of multi-level structure for improving accuracy.

As we can see in Table 3, the introduction of visual features and character-level embedding vectors can effectively improve the quality of the task. In this regard, we analyze that image information can enhance the semantic understanding of MLNet, and this multimodal BERT structure we use can filter the visual information, retain the regions with higher similarity to the text, and achieve image-text and feature fusion to avoid directly superimposing the less relevant image information on top of the text, which causes unnecessary errors. The introduction of character-level embedding is another important reason for achieving the MLNet effect. Due to the short text length of tweets, it is difficult to obtain the semantics according to the context, and there are certain misspellings, URL addresses, and some emojis, which bring interference with label classification. Preserving character-level embedding can obtain more semantic information to a certain extent and improve the effectiveness of short text recognition tasks.

4.4.2. Cross-task scenario

We further tested the performance of UMGF and MLNet in cross-domain scenarios and compared them. We used the model obtained by training on the Twitter2015 dataset to test Twitter2017 and notated as Twitter2015→Twitter2017. Similarly, Twitter2017→Twitter2015 indicates the use of a model trained on Twitter2017 to test Twitter2015. As shown in Table 4, our MLNet achieves better results in terms of F1 Score in this cross-task scenario experiment. This proves that our MLNet has made some progress in model liability.

Although MLNet is slightly less effective on the Twitter2015 dataset than the Twitter2017 dataset, our model trained on the Twitter2015 dataset is still more effective than the model trained on the Twitter2017 dataset in the migration experiment. This also shows that although our models have better results on small datasets, training MLNet on larger amounts of data is still effective and can improve the understanding of the models. This cross-migration scenario is interesting to facilitate the entity recognition task and better improve the language model.

5. Conclusion

In this study, we propose a new multilevel multichannel fusion network for the named entity recognition task in social media postings. Specifically, we propose a focused visual feature preprocessing method in multimodal tasks to extract visual features related to text semantics as auxiliary inputs, which is an effective visual enhancement of the NER attention module. We also propose

TABLE 4 Performance comparison of MLNet and UMGF in cross-task scenario.

Methods	Twitter2015→Twitter2017			Twitter2017→Twitter2015		
	Precision	Recall	F1	Precision	Recall	F1
UMGF	67.00	**62.81**	66.21	69.88	56.92	62.74
MLNet	**70.60**	62.46	**66.28**	**71.01**	**58.40**	**64.09**

The best results in the table are highlighted in bold.

the inclusion of character embedding, which expands the feature information that can be extracted from short texts and implements an entity extraction architecture with multiple levels of input. Achieving a better extractor through effective visual enhancement, extensive experiments, and results on three criteria demonstrate the effectiveness and robustness of our proposed approach. At the same time, our method faces the limitation of slower operations. In future, we plan to (1) further investigate the simplified structure of multimodal multilevel entity extraction models to make them more flexible and scalable and (2) try more diverse image-text feature fusion algorithms to help the models better understand the association between visual features and text features.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HZ and ZH contributed to conception and design of the study. XT organized the database. XL funded acquisition. HZ, ZH, and XT performed the statistical analysis. HZ wrote the first draft of the manuscript. ZH and FB wrote the sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China-China State Railway Group Co., Ltd. Railway Basic Research Joint Fund (Grant No. U2268217) and the Scientific Funding for China Academy of Railway Sciences Corporation Limited (No. 2021YJ183). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Conflict of interest

XL is employed by China Academy of Railway Sciences Corporation Limited.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Ahn, H., Choi, S., Kim, N., Cha, G., and Oh, S. (2018). Interactive text2pickup networks for natural language-based human–robot collaboration. *IEEE Robot. Automat. Lett.* 3, 3308–3315. doi: 10.1109/LRA.2018.2852786

Arshad, O., Gallo, I., Nawaz, S., and Calefati, A. (2019). "Aiding intra-text1 representations with visual context for multimodal named entity recognition," in *2019 International Conference on Document Analysis and Recognition (ICDAR)* (Sydney: International Convention Centre Sydney), 337–342.

Asgari-Chenaghlu, M., Feizi-Derakhshi, M. R., Farzinvash, L., Balafar, M., and Motamed, C. (2020). A multimodal deep learning approach for named entity recognition from social media. *arXiv [Preprint]. arXiv:2001.06888*. doi: 10.1007/s00521-021-06488-4

Bouthillier, X., Konda, K., Vincent, P., and Memisevic, R. (2015). Dropout as data augmentation. *arXiv [Preprint]. arXiv:1506.08700*. doi: 10.48550/arXiv.1506.08700

Chen, D., Li, Z., Gu, B., and Chen, Z. (2021). "Multimodal named entity recognition with image attributes and image knowledge," in *Database Systems for Advanced Applications: 26th International Conference, DASFAA 2021* (Taipei: Springer), 186–201.

Chen, S., Aguilar, G., Neves, L., and Solorio, T. (2020). Can images help recognize entities? a study of the role of images for multimodal NER. *arXiv [Preprint]. arXiv:2010.12712*. doi: 10.18653/v1/2021.wnut-1.11

Chen, X., Li, L., Deng, S., Tan, C., Xu, C., Huang, F., et al. (2021). Lightner: a lightweight tuning paradigm for low-resource ner via pluggable prompting. *arXiv [Preprint]. arXiv:2109.00720*.

Chen, X., Zhang, N., Li, L., Yao, Y., Deng, S., Tan, C., et al. (2022). Good visual guidance makes a better extractor: hierarchical visual prefix for multimodal entity and relation extraction. *arXiv [Preprint]. arXiv:2205.03521*. doi: 10.18653/v1/2022.findings-naacl.121

Chen, Y.-C., Li, L., Yu, L., El Kholy, A., Ahmed, F., Gan, Z., et al. (2020). "UNITER: universal image-text representation learning," in *Computer Vision–ECCV 2020: 16th European Conference* (Glasgow: Springer), 104–120.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. *arXiv [Preprint]. arXiv:1810.04805*. doi: 10.48550/arXiv.1810.04805

Gan, Z., Chen, Y.-C., Li, L., Zhu, C., Cheng, Y., and Liu, J. (2020). "Large-scale adversarial training for vision-and-language representation learning," in *Advances in Neural Information Processing Systems 33*, 6616–6628.

Gao, D., Jin, L., Chen, B., Qiu, M., Li, P., Wei, Y., et al. (2020). "FashionBERT: text and image matching with adaptive loss for cross-modal retrieval," in *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2251–2260.

Hammerton, J. (2003). "Named entity recognition with long short-term memory," in *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003* (Edmonton, AB), 172–175.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Las Vegas, NV), 770–778.

Hong, Y., Wu, Q., Qi, Y., Rodriguez-Opazo, C., and Gould, S. (2021). "VLN BERT: a recurrent vision-and-language bert for navigation," in *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition* (Nashville, TN), 1643–1653.

Jawahar, G., Sagot, B., and Seddah, D. (2019). "What does bert learn about the structure of language?" in *ACL 2019-57th Annual Meeting of the Association for Computational Linguistics* (Florence).

Jiang, H., Zhang, D., Cao, T., Yin, B., and Zhao, T. (2021). Named entity recognition with small strongly labeled and large weakly labeled data. *arXiv [Preprint]. arXiv:2106.08977*. doi: 10.18653/v1/2021.acl-long.140

Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., and Dyer, C. (2016). Neural architectures for named entity recognition. *arXiv [Preprint]. arXiv:1603.01360*. doi: 10.48550/arXiv.1603.01360

Li, G., Duan, N., Fang, Y., Gong, M., and Jiang, D. (2020). "Unicoder-VL: a universal encoder for vision and language by cross-modal pre-training," in *Proceedings of the AAAI Conference on Artificial Intelligence* (New York, NY), 11336–11344.

Li, J., Ding, H., Shang, J., McAuley, J., and Feng, Z. (2021). Weakly supervised named entity tagging with learnable logical rules. *arXiv [Preprint]. arXiv:2107.02282*. doi: 10.18653/v1/2021.acl-long.352

Li, L. H., Yatskar, M., Yin, D., Hsieh, C.-J., and Chang, K.-W. (2019). VisualBERT: a simple and performant baseline for vision and language. *arXiv [Preprint]. arXiv:1908.03557*. doi: 10.48550/arXiv.1908.03557

Li, Y., Fan, H., Hu, R., Feichtenhofer, C., and He, K. (2022). Scaling language-image pre-training via masking. *arXiv [Preprint]. arXiv:2212.00794*. doi: 10.48550/arXiv.2212.00794

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). "Microsoft COCO: common objects in context," in *Computer Vision–ECCV 2014: 13th European Conference* (Zurich: Springer), 740–755.

Liu, Y., Meng, F., Zhang, J., Xu, J., Chen, Y., and Zhou, J. (2019). GCDT: a global context enhanced deep transition architecture for sequence labeling. *arXiv [Preprint]. arXiv:1906.02437*. doi: 10.18653/v1/P19-1233

Lu, D., Neves, L., Carvalho, V., Zhang, N., and Ji, H. (2018). "Visual attention model for name tagging in multimodal social media," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics* (Melbourne, VIC: Association for Computational Linguistics), 1990–1999.

Luo, L., Yang, Z., Yang, P., Zhang, Y., Wang, L., Lin, H., et al. (2018). An attention-based biLSTM-CRF approach to document-level chemical named entity recognition. *Bioinformatics* 34, 1381–1388. doi: 10.1093/bioinformatics/btx761

Luo, Y., Xiao, F., and Zhao, H. (2020). "Hierarchical contextualized representation for named entity recognition," in *Proceedings of the AAAI Conference on Artificial Intelligence* (New York, NY: AAAI), 8441–8448.

Moon, S., Neves, L., and Carvalho, V. (2018). Multimodal named entity recognition for short social media posts. *arXiv [Preprint]. arXiv:1802.07862*. doi: 10.18653/v1/N18-1078

Murahari, V., Batra, D., Parikh, D., and Das, A. (2020). "Large-scale pretraining for visual dialog: A simple state-of-the-art baseline," in *Computer Vision–ECCV 2020: 16th European Conference* (Glasgow: Springer), 336–352.

Park, J. S., Jia, B., Bansal, M., and Manocha, D. (2019). "Efficient generation of motion plans from attribute-based natural language instructions using dynamic constraint mapping," in *2019 International Conference on Robotics and Automation (ICRA)* (Montreal, QC: IEEE), 6964–6971.

Pinheiro, P., and Collobert, R. (2014). "Recurrent convolutional neural networks for scene labeling," in *International Conference on Machine Learning* (Beijing: PMLR), 82–90.

Radmard, P., Fathullah, Y., and Lipani, A. (2021). "Subsequence based deep active learning for named entity recognition," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 4310–4321.

Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv [Preprint]. arXiv:1409.1556*. doi: 10.48550/arXiv.1409.1556

Souza, F., Nogueira, R., and Lotufo, R. (2019). Portuguese named entity recognition using bert-crf. *arXiv [Preprint]. arXiv:1909.10649*. doi: 10.48550/arXiv.1909.10649

Su, W., Zhu, X., Cao, Y., Li, B., Lu, L., Wei, F., and Dai, J. (2019). VL-BERT: pre-training of generic visual-linguistic representations. *arXiv [Preprint]. arXiv:1908.08530*. doi: 10.48550/arXiv.1908.08530

Tan, H., and Bansal, M. (2019). LXMERT: learning cross-modality encoder representations from transformers. *arXiv [Preprint]. arXiv:1908.07490*. doi: 10.18653/v1/D19-1514

Tian, Y., Sun, X., Yu, H., Li, Y., and Fu, K. (2021). Hierarchical self-adaptation network for multimodal named entity recognition in social media. *Neurocomputing* 439, 12–21. doi: 10.1016/j.neucom.2021.01.060

Walker, N., Peng, Y.-T., and Cakmak, M. (2019). "Neural semantic parsing with anonymization for command understanding in general-purpose service robots," in *RoboCup 2019: Robot World Cup XXIII* (Sydney, NSW: Springer), 337–350.

Wang, S., Li, X., Meng, Y., Zhang, T., Ouyang, R., Li, J., et al. (2022). *k* nn-ner: named entity recognition with nearest neighbor search. *arXiv [Preprint]. arXiv:2203.17103*. doi: 10.48550/arXiv.2203.17103

Wang, X., Jiang, Y., Bach, N., Wang, T., Huang, Z., Huang, F., et al. (2021). Improving named entity recognition by external context retrieving and cooperative learning. *arXiv [Preprint]. arXiv:2105.03654*. doi: 10.48550/arXiv.2105.03654

Wu, Z., Zheng, C., Cai, Y., Chen, J., Leung, H.-F., and Li, Q. (2020). "Multimodal representation with embedded visual guiding objects for named entity recognition in social media posts," in *Proceedings of the 28th ACM International Conference on Multimedia* (Seattle, WA), 1038–1046.

Yu, J., Jiang, J., Yang, L., and Xia, R. (2020). *Improving Multimodal Named Entity Recognition via Entity Span Detection With Unified Multimodal Transformer*. Association for Computational Linguistics.

Zhang, D., Wei, S., Li, S., Wu, H., Zhu, Q., and Zhou, G. (2021). "Multi-modal graph fusion for named entity recognition with targeted visual guidance," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 14347–14355.

Zhang, Q., Fu, J., Liu, X., and Huang, X. (2018). "Adaptive co-attention network for named entity recognition in tweets," in *Proceedings of the AAAI Conference on Artificial Intelligence* (New Orleans, LA: AAAI).

frontiers | Frontiers in Robotics and AI

TYPE Perspective
PUBLISHED 17 August 2023
DOI 10.3389/frobt.2023.1126033

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Per Anders Nilsson,
University of Gothenburg, Sweden

*CORRESPONDENCE
Kevin J. Ryan Jr.,
kryan30@utk.edu

RECEIVED 17 December 2022
ACCEPTED 01 August 2023
PUBLISHED 17 August 2023

CITATION
Ryan KJ Jr. (2023), Cognitive modeling, ecological psychology, and musical improvisation.
Front. Robot. AI 10:1126033.
doi: 10.3389/frobt.2023.1126033

Cognitive modeling, ecological psychology, and musical improvisation

Kevin J. Ryan Jr.*

Department of Philosophy, University of Tennessee, Knoxville, TN, United States

Understanding novelty and improvisation in music requires gathering insight from a variety of disciplines. One fruitful path for synthesizing these insights is via modeling. As such, my aim in this paper is to start building a bridge between traditional cognitive models and contemporary embodied and ecological approaches to cognitive science. To achieve this task, I offer a perspective on a model that would combine elements of ecological psychology (especially affordances) and the Learning Intelligent Decision Agent (LIDA) cognitive architecture. Jeff Pressing's cognitive model of musical improvisation will also be a central link between these elements. While some overlap between these three areas already exists, there are several points of tension between them, notably concerning the nature of perception and the function of artificial general intelligence modeling. I thus aim to alleviate the most worrisome concerns here, introduce several future research questions, and conclude with several points on how my account is part of a general theory, rather than merely a redescription of existent work.

KEYWORDS

improvisation, cognitive model, ecological psychology, learning intelligent decision agent (LIDA), affordances, music

1 Introduction

To understand novelty in cognition, we must account for improvisation; as philosopher Gilbert Ryle noted about improvising, "It is part of intelligence to seize new opportunities and to face new hazards; to be, in short, 'not a tram, but a bus'" [(Ryle, 1976), 69]. Other philosophers have explored improvisation in detail, including, *inter alia*, its connection to composition and repetition, possible moral dimensions, and its relation to creativity and novelty (Brown, 1996; Gould and Keaton, 2000; Alperson, 2010; Carvalho, 2010; Hagberg, 2016; Brown, Goldblatt and Gracyk, 2018; Lewis, 2019). My focus will hereafter be on musical improvisation.

The scientific study of improvisation in music has seen major development in the past several decades. Landmark work by musician and psychologist Pressing. (1988), Pressing (1998) has left a lasting impact on cognitive accounts of musical improvisation. Another important computational account was developed by psychologist and philosopher Johnson-Laird (2002). Similarly, work in artificial intelligence has been steadily advancing. Examples of notable AI programs here include Voyager (Lewis, 2000), Magenta, and BachBot (Novelli and Proksch, 2022).

Since the early 2000s, neuroscientists have discovered important roles for different neural areas in improvisation, including (but not limited to) the left inferior frontal gyrus (IFG) and the dorsolateral prefrontal cortex (DLPFC) (For a review, see Beaty, 2015; for a conceptual

model involving neural correlates, see Faber and McIntosh, 2019). Evidence has emerged for differences during solo vs. group improvisation, most notably in activations of the DLPFC. Interpretations for why these differences occur include unique requirements for creativity and monitoring in individual or collective contexts; there may also be methodological differences in participant selection and task requirements across studies (Beaty, 2015).

Another discipline to consider is ecological psychology, wherein the environment is taken to be both an inextricable part of psychological explanation and a constitutive part of cognition. A pioneer of this approach was psychologist Gibson 1966, Gibson 1979. It has further developed to include accounts of auditory cognition (Gaver, 1993a; Gaver, 1993b), music perception (Clarke, 2005), music in everyday life (DeNora, 2000, esp. Chap. 4), musical affordances (Reybrouck, 2012; Windsor and de Balzac, 2012; Krueger, 2014), and the structure of performance spaces (Burland and Windsor, 2014), among other topics. Cognitive scientist and philosopher Ashley Walton and her colleagues (Walton et al., 2015; Walton et al., 2018) have also provided similar work on dynamics and interaction in performance.

In what follows, I propose an important way to further our understanding of musical improvisation by bringing together ecological psychology, the Learning Intelligent Decision Agent (LIDA) cognitive architecture, and Pressing's model of musical improvisation. I first introduce Pressing's model. Second, I introduce core aspects of LIDA. Third, I present an ecological description of Pressing's model from music theorist and lawyer Love. (2017) and, finally, I discuss several main issues about how Pressing, Love, LIDA can be connected. While there are insights for many different readers in what follows, I hope that it speaks clearly to those who are interested with an account of improvisation in cognition and action, including the similarities and differences between domain general and domain specific aspects of improvised activity.

2 Pressing's model of musical improvisation

There is an expansive and growing literature for models, including theories and interdisciplinary work, of musical performance (e.g., see MacDonald and Wilson, 2020, Chap. 2, esp. 30–43, for a theoretical overview). There are likewise notable differences among traditions of cognitive modeling. For instance, classical approaches focus on the development of music starting within an agent, post-human approaches place emphasis on the flow of information from the environment into the agent, computational accounts take most, or all, of human cognition as akin to functions on computers, and embodied approaches center cognition as a dynamic dance between brain, body, and world.

Pressing had a classical approach to modeling musical improvisation[1]. The heart of his model is the referent and the knowledge base of an improviser. A referent is "a set of cognitive, perceptual, or emotional structures (constraints) that guide and aid in the production of musical materials" (Pressing, 1998, 52). Paradigmatic examples include a jazz standard or a 12-bar blues progression. Some referents are stored in external formats, utilizing musical notation and sheet music, while others are internalized. The purpose and existence of a referent will vary depending on the type of improvisational activity and other relevant considerations.

The knowledge base covers a range of information stored primarily in long-term memory. Specific elements include "musical materials and excerpts, repertoire, subskills, perceptual strategies, problem-solving routines, hierarchical memory structures and schemas, generalized motor programs, and more" (Pressing, 1998, 53). The base can further be connected with three mental representations: objects, features, and processes. Objects are specific "cognitive units," such as a chord or gesture; features are "common parameters of multiple objects; " processes are "changes in objects or features over time; " and, finally, it is essential to highlight that all of these aspects interact with each other in complex ways (Dean and Bailes, 2014, 41–42).

According to Pressing, improvisors execute plans to either continue current musical events via association or break via interruption. The general forms of these choices are consistent–association maintains most/all of the aspects (e.g., movement, musical, acoustic, or other features) from previous events, while interruption breaks from them to some significant degree–but their exact functioning in performance will vary based on multiple factors. For example, a bassist may decide to play a pedal point across several measures. A continuation of the tone would be an association between the starting event (E_i) and E_{i+1}, E_{i+2}, … E_{i+n}, while abruptly stopping it would be an interruption between these events. Reasons for choosing continuity or interruption include acoustic, social, musical, and movement-based considerations, among others. These processes are also facilitated by an interruption tester monitoring E_i and a movement trigger between events (For a visual overview of Pressings model, see Fig. 7.4 in Pressing, 1988, 160).

There are two further things to highlight about Pressing's model. First, it operates under an assumption that improvisation develops along a discrete repetition of input-process-output cycles. These cycles are musical events and event clusters. For instance, a given solo for a jazz bebop performance would be an event cluster constituted by multiple interlocking musical events (e.g., specific notes and changing chords). Second, his model is developed at a rather general level. In Pressing's words, for instance, the model "does not spell out exactly how real-time constraints of memory and attention are to be accommodated" (1998, 56). Connecting this model to LIDA will help to address some of these downsides and further strengthen the upsides.

3 Learning intelligent decision agent (LIDA) cognitive architecture

Learning Intelligent Decision Agent (LIDA) is a conceptual and partially computational systems-level cognitive model that aims to model mind (see Figure 1). Insofar as it operates at the systems-

1 It is important to note that Pressing's model is best fit for "standard" jazz improvisation, especially in the context of either solo performance and/or group improvisations with a clear referent. While it is not limited to these contexts, we may need to consider different models (e.g., MacDonald and Wilson, 2020, esp. Chap. 4; Canonne and Garnier, 2011) for capturing "free" jazz performance or improvisation in other genres of music.

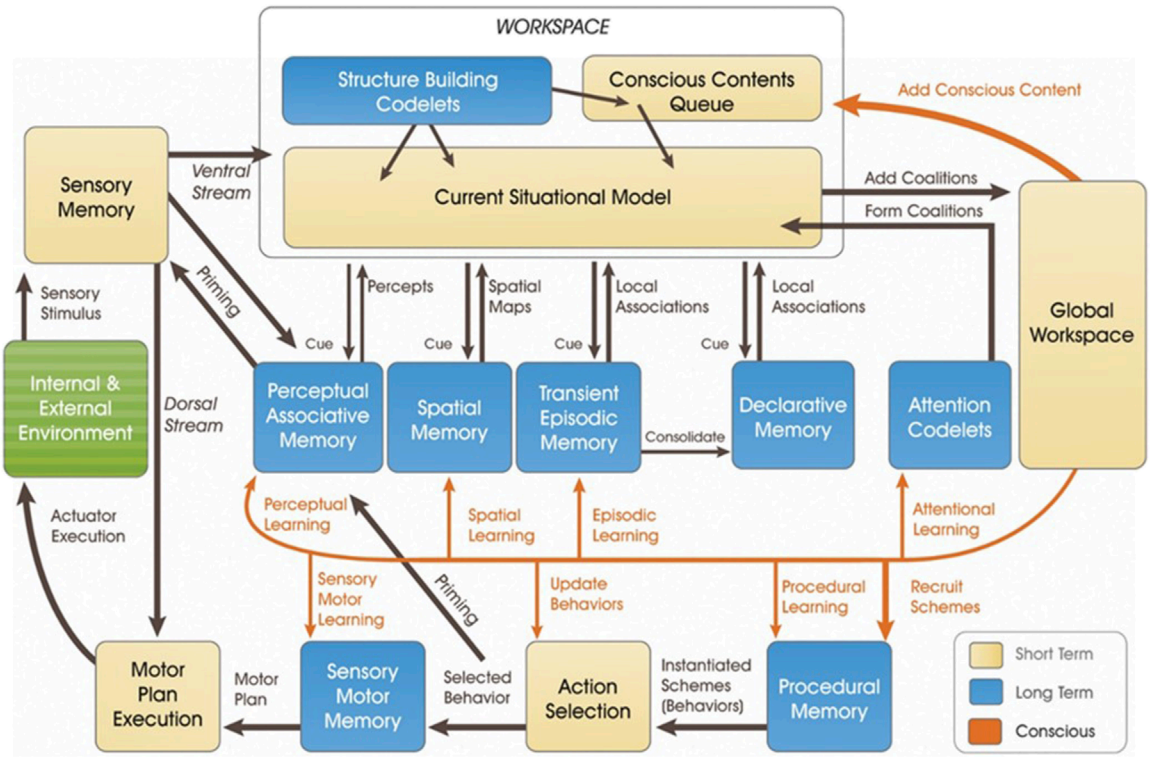


FIGURE 1
The LIDA Model Overview.

level, LIDA attempts to account for the entire span of low-level to high-level cognition. According to Stan Franklin, computer scientist and the main architect behind LIDA, a mind is a control structure of an autonomous agent (AA). AA has a technical definition of "a system situated in and part of an environment, which senses that environment and acts on it over time in accordance with its own agenda, so as it may affect what it senses in the future" (Franklin and Graesser, 1997).

A central feature of LIDA is a "cognitive cycle" wherein learning, perception, and action occur. The core part of the cognitive cycle is a global broadcast of information across the mind. The first step in the cycle is input, either exogenous (e.g., sensory input) or endogenous (e.g., memory), that results in perception and/or understanding of the current situation. Second, in the attention phase, information that has reached a certain level of salience is broadcast. Third, this broadcasted information results in learning and/or action.

Cognitive cycles can overlap and it is possible for actions to occur both asynchronously and without rising to the level of consciousness (i.e., without being globally broadcast). While the majority of LIDA research thus far has focused on fleshing out the cognitive cycle, recent developments have included cases that require multiple cognitive cycles, including distal intentions and narratives (Kronsted et al., 2021), the body schema (Neemeh et al., 2021), and smooth coping (Kronsted et al., 2022). Modeling musical improvisation will likewise require multiple cognitive cycles.

Fully implementing Pressing's model in LIDA is outside the scope of this current work, but three points can be said about a partial implementation. First, modeling expertise in musical improvisation will require further development of codelets. Codelets, especially for attention and structure-building, are important parts of LIDA. They fulfill tasks such as bringing together different types of information or raising relevant information to the level of consciousness. In addition, work on LIDA has considered a specific role for some structure-building codelets in creating affordance links (Franklin et al., 2016; Section 5.3.2). Second, further developing the Perceptual Associative Memory and Procedural Memory modules will be important for modeling skills in musical improvisation, especially involving a referent. Third, the Conscious Contents Queue, Action Selection and Motor Plan Execution modules will be essential parts of grounding any model of improvisation. Part of developing these modules will connect to how improvisation occurs with and without consciousness. Figure 2.

One may wonder why I appeal to LIDA, rather than another Artificial General Intelligence (AGI) model. On one hand, it is useful to develop as many cognitive architectures as possible. On the other, LIDA offers several unique aspects that have been underexplored in AGI. For example, LIDA's focus on AAs means that it does not reduce improvisation to a problem-solving task, since the explorations, sensations, and alterations that are part of being an AA covers a wide swath of why we improvise. The central focus on affect, emotion, and consciousness also makes LIDA an important architecture to use and further develop; indeed, while affect and emotion are an essential part of all human cognition, they play an extremely strong role

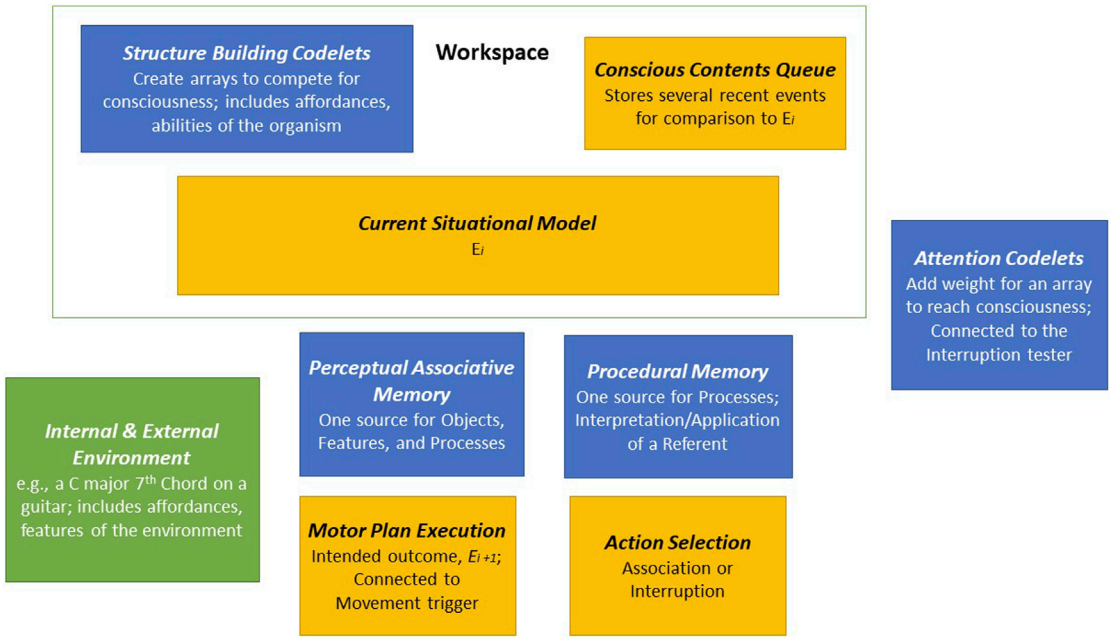


FIGURE 2
Musical Improvisation in LIDA–A Partial Conceptual Implementation.

in music (Schiavio et al., 2017; Van der Schyff and Schiavio, 2017; Novelli and Proksch, 2022).

4 An ecological description and musical affordances in improvisation

Love (2017) provides an ecological description of jazz improvisation. His paper begins with a concern of oversimplification for computational models. Specifically, while simplification is needed for computational purposes, it does not match actual human improvising (Love, 2017, 32). He raises a connected worry about any neat distinctions between input, processing, and output, since cognition in the world is neither that simple nor clean. In turn, Love develops his description by reformulating ideas from Pressing's model, including the referent, the role of memory and learning, and the temporal coordination among soloists and ensembles. I shall focus on the referent and memory/learning in what follows.

First, according to Pressing, a benefit of the referent is to simplify improvisation. Following chord changes allows a soloist to focus on other aspects of creativity, for instance. Love grants this simplification as important. However, he notes that a referent may also make improvising more difficult "by introducing the possibility of failure, or, if we prefer, shrinking the set of actions that count as 'success'" (Love, 2017, 34). This difficulty is connected to the idea that improvisation is a form of navigation across a terrain, rather than abstract problem-solving. Any solo navigational difficulty is moreover amplified within collective forms of musical improvisation, where ensembles support or challenge the soloist navigating their environment (Linson and Clarke, 2017).

Second, instead of long-term memory, Love's ecological description focuses on perceptual process and affordances. An affordance is a relational property between the abilities of an organism and features of the environment (Chemero, 2009). For instance, the affordance "climbable" is present if an organism is able to scale vertical surfaces in the environment; a specific "climbable" surface for a human may not exist for a dog, just as a professional climber may find surfaces "climbable" that are impossible for a novice. There is also an important role for affordances in the design of everyday objects, as emphasized by engineer and design expert Norman (2013).

Following philosophers Rietveld and Kiverstein. (2014), I suggest that affordances are found in a landscape for a form of life. Any situation will include a salient field of affordances within the overall landscape (Einarsson and Ziemke, 2017). The form of life is the focal point since it helps make sense for how affordances come into existence and how they persist over time.

Finally, Love cites several additional pieces of evidence about memory, including worries about the overemphasis of memory across perception and the fact that much of what an improvisor does is connected to how they perceive and navigate an improvisation, instead of using an abstract store of representations separate from perception itself (Love, 2017, 38–40).

I have been calling Love's work a description, rather than a theory or model. The reason why, as Love himself notes, is his account lacks core aspects of a theory, especially falsifiability (Love, 2017, 31). By connecting Pressing and Love directly to LIDA, I suggest that we will at least take steps in the direction of a developed theory, even though I do not provide that fully developed theory here. Furthermore, LIDA is integrally connected to several additional theoretical and empirical commitments, most notably

Bernard Baars' Global Workspace Theory (GWT) (Franklin et al., 2016; Section 4.4). These connections help support the idea that my proposed account will be part of a general theory, not merely a description.

5 Discussion

There are two main points I will consider in this discussion: First, does my model call for a major reformulation or minor refocus in Pressing's account? Second, is LIDA able to fully capture the affordances that are essential to an ecological account of improvisation?

For the first point, I suggest that the answer will partially rest on the extent to which Pressing's model contains the ecological aspects highlighted by Love. At first glance, Pressing's model already includes an explicit role for performer-environmental interaction. He likewise emphasizes a refinement of perception and action in musical performance, in part, by learning to discern invariant structures of the environment. Pressing introduces invariants in ways that are similar to affordances. When discussing "arrays" of objects, features, and processes, for instance, he notes that "the answer given here is based on an ecological perspective, which considers that the capacity to extract or create such arrays is neurologically innate, but that they are only brought into being by interaction with the environment" (Pressing, 1988, 161). We find a similar situation with other parts of the knowledge base, notably "perceptual strategies" (1998, 53), which implies that perception does not collapse into another form of memory.

The two possible results to the first point are either (1) Love is offering a minor shift in emphasis rather than a major reformulation or (2) Love's reformulation comes mainly from worries that Pressing's account fails to capture other essential aspects of affordances and/or the environment beyond invariance. While I am not committed to either of these interpretations, the best likelihood for (2) would come from the subjective component of affordances and their "ambivalent relationship to rules" (Love, 2017, 34). (2) may also be supported by Pressing's explicit rejection of "the organizational invariant approach" as a satisfactory theory for modeling action and improvisation (1988, 133–4).

This rejection of an organizational invariant approach raises a concern that Pressing's account is not about the perception of affordances because it requires a comparison of sensation to pre-existing knowledge, instead of the agent adapting to affordances in the environment. One reply to this concern is to grant that Pressing's approach and affordances are more opposed than I heretofore implied, with substantial work required to reconcile them, if doing so is even possible. A second reply is to consider recent work on neural resonance as a way to account for organism abilities and the knowledge base in an ecological manner (Fuchs, 2018; Raja, 2018; Ryan and Gallagher, 2020; Shepard, 1984; for neural resonance and neurodynamics in musical perception, see Large, 2010). This second option, if correct, still requires more work, but it would be closer to a slight refocus over radical reformulation.

For the second main point, one may be concerned with the vast assortment of memory modules in LIDA with limited cases of distinctly perceptual processes. Another issue may be that the affordances introduced so far in LIDA are not adequately relational. To address these worries, I will say a few words on the connection between LIDA and embodied cognition, along with the importance of looking at the implementation of the LIDA model in LIDA agents as part of the modeling process.

Franklin et al. (2016) argue that LIDA is "resonant with the core ideas of the embodied, situated, and enactive views" (2016, Section 4.2), with situated cognition being close to the core commitments of ecological psychology introduced so far (i.e., the mutual influence and interaction of environment and organism in cognition). Three points of connection they discuss in detail are asynchrony, nonlinear dynamics, and Theta Gamma Coupling. Of primary note here is asynchrony. While much of the description of LIDA introduces the model as if it works in a serial fashion, that is only for ease of explanation. In practice, almost all of the processing in the LIDA architecture may occur in asynchronous fashion, which entails that the "LIDA model accommodates the possibility of algorithmic behavior more complex than that of a data pipeline in the information processing paradigm." (Franklin et al., 2016 Section 4.9). Breaking down this pipeline is akin to what Love calls for when he questions the neat distinction of input-process-output.

Additionally, an implementation of LIDA in a LIDA agent is essential when considering the nature of affordances themselves. An affordance is a relational property between an organism and the environment. While certain abstract claims can be made about them, an affordance must be grounded in specific organisms acting in specific environments. Similarly, affordances in LIDA are partially constituted by implementation in an agent, rather than being fully, and solely, developed in the conceptual model. It may even be the case that using the model alone to tell us everything about affordances could be a category mistake.

6 Conclusion

I have proposed that a fruitful avenue for understanding musical improvisation will come from a combination of Pressing's model, taken under Love's ecological description, and the Learning Intelligent Decision Agent (LIDA) cognitive architecture. My theory includes at least the following three points, which are open to further exploration, refinement, or change as additional research is conducted.

- Insights from a traditional cognitive model and an ecological description of that model can be fruitfully combined within an AGI architecture;
- While this combination draws on existent theoretical components, it also opens up important theoretical and empirical questions for the future (e.g., how can we build an improvising LIDA robot? In what ways will it be unique and similar to existent improvising AI machines? How can we use the LIDA cognitive model and musical improvisation to test open questions about Global Workspace Theory and other accounts of consciousness? What aspects of improvisation occur within and outside of consciousness?);
- An account of musical improvisation is integrally connected to other forms of improvisation, and we cannot fully understand any of those activities in complete isolation.

Moving forward, it will likely be important to further nest the aforementioned approaches within even broader theoretical frameworks, such as the Skilled Intentionality Framework (Rietveld, Denys and Van Westen, 2018) or the Thinking Through Other Minds and Cultural Affordances Frameworks (Ramstead, Veissière and Kirmayer, 2016; Veissière et al., 2020). Integrating more interdisciplinary sources and artistic research will be essential as well, including work by Derek Bailey, Gary Peters, Marcel Cobussen, and Harald Stenström (e.g., Bailey, 1992; Stenström, 2009; Cobussen, 2017; Peters, 2017). Such additions will further develop this model of musical improvisation and its related upshots.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

KR wrote the manuscript, contributed to manuscript revision, read, and approved the submitted version.

Acknowledgments

The author thanks an reviewer for their insightful comments and helpful feedback on this paper. Also, thanks to Sheri Markose and the Frontiers Editorial and Production Teams for their help throughout the peer review and publication process.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alperson, P. (2010). A topography of improvisation. *J. Aesthet. Art Crit.* 68, 273–279. doi:10.1111/j.1540-6245.2010.01418.x

Bailey, D. (1992). *Improvisation: Its nature and practice in music*. Boston, MA: De Capo Press.

Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neurosci. Biobehav. Rev.* 51, 108–117. doi:10.1016/j.neubiorev.2015.01.004

Brown, L. B., Goldblatt, D., and Gracyk, T. (2018). *Jazz and the philosophy of art*. New York, NY: Routledge.

Brown, L. B. (1996). Musical works, improvisation, and the principle of continuity. *J. Aesthet. Art Crit.* 54, 353–369. doi:10.2307/431917

Burland, K., and Windsor, W. L. (2014). "Moving the gong: exploring the contexts of improvisation and composition," in *Coughing and clapping: Investigating audience experience*. Editors K. Burland, and S. Pitts (London, United Kingdom: Routledge).

Canonne, C., and Garnier, N. (2011). "A model for collective free improvisation," in *Mathematics and computation in music. MCM 2011*. Editors C. Agon, M. Andreatta, G. Assayag, E. Amiot, J. Bresson, and J. Mandereau (Berlin, Heidelberg: Springer). doi:10.1007/978-3-642-21590-2_3

Carvalho, J. M. (2010). Repetition and self-realization in jazz improvisation. *J. Aesthet. Art Crit.* 68, 285–290. doi:10.1111/j.1540-6245.2010.01420.x

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Clarke, E. F. (2005). *Ways of listening: An ecological approach to perception of musical meaning*. Oxford, UK: Oxford U. Press.

Cobussen, M. (2017). *The field of musical improvisation*. Leiden, NL: Leiden U. Press.

Dean, R. T., and Bailes, F. (2014). "Cognitive processes in musical improvisation," in *The oxford handbook of critical improvisation studies*. Editors G. E. Lewis, and B. Piekut (Oxford, UK: Oxford U. Press).

DeNora, T. (2000). *Music in everyday life*. Cambridge, UK: Cambridge U. Press.

Einarsson, A., and Ziemke, T. (2017). Exploring the multi-layered affordances of composing and performing interactive music with responsive technologies. *Front. Psychol.* 8, 1701. doi:10.3389/fpsyg.2017.01701

Faber, S. E. M., and McIntosh, A. R. (2019). Towards a standard model of musical improvisation. *Eur. J. Neurosci.* 51, 840–849. doi:10.1111/ejn.14567

Franklin, S., and Graesser, A. C. (1997). "Is it an agent, or just a program? A taxonomy for autonomous agents," in *Intelligent agents III (21–35)* (Berlin: Springer Verlag).

Franklin, S., Madl, T., Strain, S., Faghihi, U., Dong, D., Kugele, S., et al. (2016). A LIDA cognitive model tutorial. *Biol. Inspired Cogn.* 16, 105–130. doi:10.1016/j.bica.2016.04.003

Fuchs, T. (2018). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford: Oxford U. Press.

Gaver, W. W. (1993b). How do we hear in the world? Explorations in ecological acoustics. *Ecol. Psychol.*, 5, 285–313.

Gaver, W. W. (1993a). What in the world do we hear? An ecological approach to auditory event perception. *Ecol. Psychol.* 5, 1–29. doi:10.1207/s15326969eco0501_1

Gibson, J. J. (1979). *The ecological approach to visual perception*. Oxford, UK: Houghton Mifflin.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Oxford, UK: Houghton Mifflin.

Gould, C. S., and Keaton, K. (2000). The essential role of improvisation in musical performance. *J. Aesthet. Art Crit.* 58, 143–148. doi:10.1111/1540-6245.jaac58.2.0143

Hagberg, G. (2016). "Ensemble improvisation, collective intention, and group attention," in *The Oxford handbook of critical improvisation studies*. Editors G. E. Lewis, and B. Piekut (New York, NY: Oxford U. Press), 481–499.

Johnson-Laird, P. N. (2002). How jazz musicians improvise. *Music Percept.* 19, 415–442. doi:10.1525/mp.2002.19.3.415

Kronsted, C., Kugele, S., Neemeh, Z. A., Ryan, K., and Franklin, S. (2022). Embodied intelligence: smooth coping in the learning intelligent decision agent (LIDA) cognitive architecture. *Front. Psychol.* 13, 846931. doi:10.3389/fpsyg.2022.846931

Kronsted, C., Neemeh, Z. A., Kugele, S., and Franklin, S. (2021). Modeling long-term intentions and narratives in autonomous agents and narratives in autonomous agents. *J. Artif. Intell. Conscious.* 8, 229–265. doi:10.1142/s2705078521500107

Krueger, J. (2014). Affordances and the musically extended mind. *Front. Psychol.* 4, 1003–1013. doi:10.3389/fpsyg.2013.01003

Large, E. W. (2010). "Neurodynamics of music," in *Springer handbook of auditory research: Music perception*. Editors M. Riess Jones, R. R. Fay, and A. N. Popper (New York, NY: Springer), 201–231. doi:10.1007/978-1-4419-6114-3_7

Lewis, E. (2019). *Intents and purposes: Philosophy and the aesthetics of improvisation*. Ann Arbor: U. of Michigan Press.

Lewis, G. E. (2000). Too many notes: computers, complexity and culture in voyager. *Leonardo music J.* 10, 33–39. doi:10.1162/096112100570585

Linson, A., and Clarke, E. F. (2017). "Distributed cognition, ecological theory and group improvisation," in *Distributed Creativity: Collaboration and Improvisation in Contemporary Music*. Editors E. F. Clarke, and M. Doffman (New York, NY: Oxford U. Press), 52–69.

Love, S. C. (2017). An ecological description of jazz improvisation. *Psychomusicology Music, Mind, Brain* 27, 31–44. doi:10.1037/pmu0000173

Macdonald, R. A. R., and Wilson, G. B. (2020). *The art of becoming: How group improvisation works*. Oxford, UK: Oxford U. Press.

Neemeh, Z. A., Kronsted, C., Kugele, S., and Franklin, S. (2021). Body schema in autonomous agents agents. *J. Artif. Intell. Conscious.* 8, 113–145. doi:10.1142/s2705078521500065

Norman, D. A. (2013). *Design of everyday things: Revised and expanded*. New York: MIT Press.

Novelli, N., and Proksch, S. (2022). Am I (deep) blue? Music-making AI and emotional awareness. *Front. Neurorobotics* 16, 897110. doi:10.3389/fnbot.2022.897110

Peters, G. (2017). *Improvising improvisation: from out of philosophy, music, dance, and literature*. Chicago: Chicago U. Press.

Pressing, J. (1988). "Improvisation: methods and a model," in *Generative processes in music: The psychology of performance, improvisation, and composition (129–178)*. Editor J. A. Sloboda (New York, NY: Oxford U. Press).

Pressing, J. (1998). "Psychological constraints on improvisational expertise and communication," in *In the course of performance: Studies in the world of musical improvisation*. Editors B. Nettl, and M. Russel (Chicago, IL: U. Chicago Press), 47–67.

Raja, V. (2018). A theory of resonance: towards an ecological cognitive architecture. *Minds Mach.* 28, 29–51. doi:10.1007/s11023-017-9431-8

Ramstead, M. J. D., Veissière, S. P. L., and Kirmayer, L. J. (2016). Cultural affordances:Scaffolding local worlds through shared intentionality and regimes of attention. *Front. Psychol.* 7, 1–21. doi:10.3389/fpsyg.2016.01090

Reybrouck, M. (2012). Musical sense-making and the concept of affordance: an ecosemiotic and experiential approach. *Biosemiotics* 5, 391–409. doi:10.1007/s12304-012-9144-6

Rietveld, E., Denys, D., and Van Westen, M. (2018). "Ecological-enactive cognition as engaging with a field of relevant affordances: the skilled intentionality framework (SIF)," in *Oxford handbook of 4e cognition*. Editors A. Newen, L. de Bruin, and S. Gallagher (Oxford: Oxford U. Press).

Rietveld, E., and Kiverstein, J. (2014). A rich landscape of affordances. *Ecol. Psychol.* 26, 325–352. doi:10.1080/10407413.2014.958035

Ryan, K. J., and Gallagher, S. (2020). Between ecological psychology and enactivism: is there resonance? *Front. Psychol.* 11, 1147. doi:10.3389/fpsyg.2020.01147

Ryle, G. (1976). Improvisation. *Mind* 85, 69–83. doi:10.1093/mind/lxxxv.337.69

Schiavio, A., van der Schyff, D., Cespedes-Guevara, J., and Reybrouck, M. (2017). Enacting musical emotions. sense-making, dynamic systems, and the embodied mind musical emotions. Sense-making, dynamic systems, and the embodied mind. *Phenomenology Cognitive Sci.* 16, 785–809. doi:10.1007/s11097-016-9477-8

Shepard, R. N. (1984). Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming perceiving, imagining, thinking, and dreaming. *Psychol. Rev.* 91, 417–447. doi:10.1037/0033-295x.91.4.417

Stenstrom, H. (2009). Free ensemble improvisation. Doctoral Dissertation. Gothenburg: Academy of Music and Drama, Faculty of Fine, Applied and Performing Arts, University of Gothenburg.

Van der Schyff, D., and Schiavio, A. (2017). The future of musical emotions. *Front. Psychol.* 8, 998. doi:10.3389/fpsyg.2017.00988

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., and Kirmayer, L. J. (2020). Thinking through other minds: a variational approach to cognition and culture. *Behav. Brain Sci.* 43, E90. doi:10.1017/s0140525x19001213

Walton, A., Richardson, M. J., Langland-Hassan, P., and Chemero, A. (2015). Improvisation and the self-organization of multiple musical bodies. *Front. Psychol.* 06, 313. doi:10.3389/fpsyg.2015.00313

Walton, A., Richardson, M. J., Langland-Hassan, P., Chemero, A., and Washburn, A. (2018). Creating time: social collaboration in music improvisation. *Top. Cognitive Sci.* 10, 95–119. doi:10.1111/tops.12306

Windsor, W. L., and de Balzac, C. (2012). Music and affordances. *Music. Sci.* 16, 102–120. doi:10.1177/1029864911435734

frontiers | Frontiers in Robotics and AI

TYPE Original Research
PUBLISHED 14 August 2024
DOI 10.3389/frobt.2024.1440631

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
José Antonio Cervantes,
University of Guadalajara, Mexico
Joshua Ljubo Bensemann,
The University of Auckland, New Zealand

*CORRESPONDENCE
Robert Johansson,
robert.johansson@psychology.su.se

RECEIVED 29 May 2024
ACCEPTED 30 July 2024
PUBLISHED 14 August 2024

CITATION
Johansson R (2024) Machine Psychology: integrating operant conditioning with the non-axiomatic reasoning system for advancing artificial general intelligence research .
Front. Robot. AI 11:1440631.
doi: 10.3389/frobt.2024.1440631

Machine Psychology: integrating operant conditioning with the non-axiomatic reasoning system for advancing artificial general intelligence research

Robert Johansson[1,2]*

[1]Department of Psychology, Stockholm University, Stockholm, Sweden, [2]Department of Computer and Information Science, Linköping University, Linköping, Sweden

This paper presents an interdisciplinary framework, Machine Psychology, which integrates principles from operant learning psychology with a particular Artificial Intelligence model, the Non-Axiomatic Reasoning System (NARS), to advance Artificial General Intelligence (AGI) research. Central to this framework is the assumption that adaptation is fundamental to both biological and artificial intelligence, and can be understood using operant conditioning principles. The study evaluates this approach through three operant learning tasks using OpenNARS for Applications (ONA): simple discrimination, changing contingencies, and conditional discrimination tasks. In the simple discrimination task, NARS demonstrated rapid learning, achieving 100% correct responses during training and testing phases. The changing contingencies task illustrated NARS's adaptability, as it successfully adjusted its behavior when task conditions were reversed. In the conditional discrimination task, NARS managed complex learning scenarios, achieving high accuracy by forming and utilizing complex hypotheses based on conditional cues. These results validate the use of operant conditioning as a framework for developing adaptive AGI systems. NARS's ability to function under conditions of insufficient knowledge and resources, combined with its sensorimotor reasoning capabilities, positions it as a robust model for AGI. The Machine Psychology framework, by implementing aspects of natural intelligence such as continuous learning and goal-driven behavior, provides a scalable and flexible approach for real-world applications. Future research should explore using enhanced NARS systems, more advanced tasks and applying this framework to diverse, complex tasks to further advance the development of human-level AI.

KEYWORDS

artificial general intelligence (AGI), operant conditioning, non-axiomatic reasoning system (NARS), machine psychology, adaptive learning

1 Introduction

Artificial General Intelligence (AGI) is the task of building computer systems that are able to understand or learn any intellectual task that a human being can. This type of AI is often contrasted with narrow or weak AI, which is designed to perform a narrow task (e.g., facial recognition or playing chess). There are several diverse research approaches

to AGI including brain-based approaches (e.g., Hawkins, 2021), projects that aims to implement different cognitive functions separately (e.g., Laird, 2019), and principle-based approaches (e.g., Hutter, 2004; Wang, 2013). Recently, Large Language Models like GPT-4 have also been introduced as a potential pathway towards achieving more generalizable AI systems (Bubeck et al., 2023).

One major challenge in contemporary AGI research is the lack of coherent theoretical frameworks (Wang, 2012; 2019). This scarcity of unified models to interpret and guide the development of AGI systems seems to have led to a fragmented landscape where researchers often work in isolation on narrowly defined problems. Coherent research frameworks could also provide a roadmap and evaluation criteria for AGI development, fostering more collaborative and interdisciplinary efforts. The fact that AGI research has not progressed as rapidly as some had hoped could very well be attributed to the lack of these comprehensive frameworks and the absence of standardized benchmarks for measuring progression towards AGI capabilities.

This work aims to address this challenge by proposing a novel framework that outlines key milestones and metrics for evaluating progress in the field of artificial general intelligence (AGI). A fundamental assumption in this work is that *adaptation* is at the heart of general intelligence. Adaptation is typically divided in *ontogenetic adaptation*, which involves the changes that occur with an organism over its lifespan, and 2) *phylogenetic adaptation*, which refers to the evolutionary changes that occur across generations within a species.

Learning has within the field of learning psychology, been equated with ontogenetic adaptation, where an individual's experiences directly impact its capabilities and behaviors (De Houwer et al., 2013). *Operant conditioning* is one type of learning that involves adaptation in the form of behavioral changes due to consequences of actions. Given the enormous amount of empirical progress generated by operant conditioning research in learning psychology, the principle of operant conditioning and its associated research tradition could be a guiding principle for AGI research.

One particular approach to building AGI is the Non-Axiomatic Reasoning System (NARS) (Wang, 2013; 2022). NARS is an adaptive reasoning system that operates on the principle of insufficient knowledge and resources, a condition that is often true for real-world scenarios. Hence, NARS is a principle-based approach that aims to address the challenges of building AGI systems that can operate effectively in dynamic and unpredictable environments (Wang, 2019). There are several NARS implementations available. One implementation is OpenNARS for Applications (ONA), that is designed to provide a practical framework for integrating NARS into various applications, with a particular focus on robotics (Hammer and Lofthouse, 2020; Hammer et al., 2023). ONA is built with sensorimotor reasoning at its core, enabling it to process sensory data in real-time and respond with appropriate motor actions. Sensorimotor reasoning, as implemented in ONA, permits the system to make sense of the world much in the same way as biological organisms do, by directly interacting with its environment and learning from these interactions. The fact that NARS systems are focused on adaptation, and that ONA has a strong emphasis on sensorimotor capabilities, suggests that they are particularly well-suited for implementing the principle of operant conditioning.

This work presents Machine Psychology, an interdisciplinary framework for advancing AGI research. It integrates principles from learning psychology, with the theory and implementation of NARS. Machine Psychology starts with the assumption that adaptation is fundamental to intelligence, both biological and artificial. As it is presented here, Machine Psychology is guided by the theoretical framework of learning psychology, and the principle of operant conditioning in particular. A sensorimotor-only version of ONA (Hammer, 2022) is used to demonstrate the feasibility of using these principles to guide the development of intelligent systems. One way to describe the integration of operant conditioning and NARS presented in this paper is that the ability to learn and adapt based on feedback from the environment, is implemented using sensorimotor reasoning that is the core of ONA. An analogy is that a neurobiological explanation of operant conditioning could be argued to be part of a biological basis for adaptive behaviors observed in many species (Brembs, 2003), similarly, the implementation within ONA using temporal and procedural inference rules offers an alternative explanation of the core of adaptive behavior and cognition.

We evaluate the Machine Psychology framework by carrying out three operant learning tasks with NARS. The first is a simple discrimination task in where NARS needs to learn, based on feedback, to choose one stimulus over another, demonstrating a fundamental aspect of learning based on the consequences of actions. The second task is more complex than the first in that the conditions of the experiment is changed midway through the task, requiring NARS to adapt its choice strategy based on the new conditions. The third experiment is a conditional discrimination task, where NARS is presented with pairs of stimuli, and must learn to select the correct stimulus based on a conditional cue that changes throughout the task, requiring an increased level of adaptability. Methods from learning psychology are used to design the experiments and guide the evaluation of the results. We explain the results by describing how the sensorimotor reasoning used by ONA enables it to adaptively modify its behavior based on the consequences of its actions. The Machine Psychology framework is demonstrated to provide a coherent experimental approach to studying the core of learning and cognition with artificial agents, and also offers a scalable and flexible framework that potentially could significantly advance research in Artificial General Intelligence (AGI).

The paper is organized as follows. Section 2 presents a background on the principles of operant conditioning and its significance in learning psychology. Section 3 introduces NARS with a focus on its foundational concepts. Section 4 describes the architecture of OpenNARS for Applications with a particular focus on its sensorimotor reasoning abilities. Section 5 discusses related work to our approach. Section 6 presents the Machine Psychology framework and how it integrates operant conditioning principles with NARS. Section 7 describes the details of the methods and experimental setup used in the evaluation of our approach. Section 8 presents the results from the experiments. Section 9 concludes the paper and outlines how the Machine Psychology framework could be used to further advance the field of AGI.

2 Operant conditioning

The work presented in this paper takes a *functional* approach to learning and adaptation (De Houwer and Hughes, 2020), and to science in general. Such approach to learning is rooted in the principles of behaviorism, which emphasizes the role of environmental interactions in shaping behavior, rather than mechanistic explanations of how internal processes affect behavior. It stems particularly from the work of B. F. Skinner, who laid much of the groundwork for understanding how consequences of an action affect the likelihood of that action being repeated in the future (Skinner, 1938). Skinner was influenced by physicist and philosopher Ernst Mach, who emphasized the use of *functional relations* in science to describe relations between events, rather than using a traditional mechanistic causal framework (Chiesa, 1994).

In the functional learning psychology tradition, learning (as ontogenetic adaptation) is defined as a change in behavior due to regularities in the environment (De Houwer et al., 2013). Several types of learning can be classified under this perspective. Operant conditioning is defined as a change in behavior due to regularities between behavior and stimuli (De Houwer et al., 2013). Other types of learning can be defined based on other the regularities in operation. A few comments regarding these definitions of learning and operant conditioning follow, as clarified by De Houwer et al. (2013). First, in line with Skinner (1938) *behavior* is defined very broadly, encompassing any observable action or response from an organism. This includes responses that are only in principle observable, such as internal physiological changes, neural processes or cognitive events. In addition to this, behavior is defined to always be a function of one or more stimuli, while a response is just an observable reaction. Second, *regularities* are defined to be any patterns of events or behavior that go beyond a single occurrence. This can be the same events happening repeatedly, or two or more events or behaviors happening at the same time. Third, the definition signals a particular view of causality ("due to"). As highlighted above, this research tradition emphasizes *functional relations* between environmental regularities and changes in behavior. This implies that learning, from this perspective, cannot be directly observed, but must be inferred from the systematic changes in behavior in response to modifications in the environment (De Houwer et al., 2013). Such inferences do depend on an observer, whose scientific goals and theoretical orientations shape the interpretation. An example follows that aims to clarify this definition further.

2.1 The three-term contingency

To describe behavior as an interaction between the organism and its environment, Skinner introduced the concept of the three-term contingency, which consists of a discriminative stimulus (Sd), a response (R), and a resultant stimulus (Sr) Skinner (1953); De Houwer and Hughes (2020). In some contexts, the terms antecedent, behavior, and consequence are used to reflect the same triadic relationship.

In many situations, an additional element is crucial: the *establishing operation* (EO), which is an example of a *motivating operation*. The EO modifies the efficacy of a resultant stimulus as a reinforcer. For instance, a consequence such as food might only be effective as a reinforcer under certain conditions, for example, after food deprivation - which would then be the establishing operation. Hence, the EO could be included as a fourth term in the three-term contingency description.

Furthermore, a *conditional discriminative stimulus* can modify the contingency based on additional contextual factors. This stimulus signals whether the relation between the discriminative stimulus, the response, and the resultant stimulus would hold or not, adding another layer of complexity to the model Lashley (1938); Stewart and McElwee (2009). The functions of the other terms would be conditional that stimulus, which explains the name conditional discriminative stimulus.

2.2 An example of operant conditioning

Imagine a rat in an experimental chamber used to study behavior. The chamber contains a small loudspeaker, a lever that can be pressed, and a water dispenser where the delivery of water is controlled by the researcher. The researcher aims to shape the rat into turning around when techno music is played and pressing the lever when classical music is played. To do this, the researcher uses an operant conditioning procedure. Before the experiment, the rat has been deprived of water for a short period, serving as the establishing operation (EO), which increases the effectiveness of water as a reinforcer.

Initially, any tendency to turn around when techno music is played (the discriminative stimulus, Sd) might be followed by the delivery of water (the resultant stimulus, Sr). This makes the behavior (R) more likely to occur in the future under similar conditions, meaning that the water functioned as a reinforcer. Conversely, when classical music (another Sd) is played, and the rat presses the lever (R), the delivery of water (Sr) follows the behavior as well. Over time, the rat learns to turn around or press the lever based on the type of music that is playing. A video of a rat performing in this experiment can be found online (WMU Rat Lab, 2009).

We could also have imagined enhancing the procedure with adding a light that could be on or off, signaling if the relation between music type, the rat's behavior, and water would hold or not. Since the function of the other terms would be conditional on the state of the light, the light would function as conditional discriminative stimulus.

This example can be considered an effective demonstration of operant conditioning. The behavior in this example is a function of both the music and the water. It illustrates the point made above that behavior studied from this perspective is not just an isolated motor action, but is also significantly influenced by the surrounding environment and the consequences that follow the behavior. The regularities in operation are reoccurring patterns of behavior and stimuli, for example, lever pressing and the delivery of water, but also a regularity regarding classical music and lever pressing. To make a causal statement about learning, we would need to observe the rat before and after interacting with these stimuli. Before the interaction, we might, for example, observe the rat exploring the cage or do random actions when different types of music was playing. After the interaction however, if we observe the behaviors described in the example, this would be a clear change in behavior

- from, for example, cage exploring when classical music is playing to lever pressing when the same music is playing. If we could argue that the change in behavior is due to the procedural arrangements (regularities), then we could potentially claim that this qualifies as an instance of learning. The type of learning it would indicate is operant conditioning since the regularities involved were between responses and stimuli (rather than, for example, a repeated pairing of stimuli as with classical conditioning). More specifically, it would be an instance of *positive reinforcement*, a kind of operant conditioning that involves an increase in target behavior due to the consequences.

2.3 Three levels of analysis

A learning situation such as the one illustrated in the example can be analyzed on three levels using learning psychology: 1) The descriptive level (or level of procedure), 2) The functional level (or the level of effect), and 3) The cognitive level (or the level of mechanism) (De Houwer and Hughes, 2020). At the descriptive level, the procedural arrangement is from the perspective of the researcher. The different sounds is used to signal if a relationship between behaving in a certain way and water holds. This is a description of procedures initiated by the researcher. It does not mean that the rat has learned based on these arrangements. The functional level however, is closer to describing the relations from the rat's perspective. If the rat turns around if and only if the techno music is playing, then that music functions as a cue for that behavior. Similarly, it is only if the delivery of water increased the behavior, that it functions as a reinforcer. If there is no change in lever pressing or turning due to the water being delivered, then that consequence has no effect. Finally, the cognitive level can be used to describe certain mental mechanisms, like association formations, that could explain how the operant learning processes take place (De Houwer and Hughes, 2020).

The importance of distinguishing these levels cannot be overstated. Just using an operant conditioning procedure (like the one above), does not mean that the subject learns in the form of operant conditioning. When doing functional learning research, we arrange procedures and study the effect on behavior change. Learning, from this perspective, is hence defined on the functional level. As stated above, a term such as reinforcement is also defined as an effect rather than a mechanism. This also means that explanations on the cognitive/mechanistic level are not part of the learning definition (De Houwer et al., 2013). It also opens up for different kinds of explanations in terms of mechanisms - for example, propositional networks as something different from association-based learning theories.

2.4 Where is the organism?

In a functional analysis of behavior, it is an interaction between organism and environment that is being analyzed. Hence, it is not the organism itself in isolation that is of interest. Technically, it is interactions between stimulus functions and response functions that are being studied, for example, an interaction between seeing a lever and pressing it, or hearing techno music and reacting to it. There are other conceptual schemas in functional learning psychology than the three-term contingency that takes into account the complexity of these interactions (Hayes and Fryling, 2018), but for this paper, what has been presented above is a sufficient conceptual framework. Importantly though, in functional learning research, these relations between procedural arrangements and behavior change does depend on an organism in that they enable response functions. This does not in an way mean that the organism causes behavior. Rather, the organism can be considered a *participant* in the arrangements (Roche and Barnes, 1997).

Recently, De Houwer and Hughes (2022) extended their conceptual work on learning beyond that of organisms, for example, to also incorporate the study of learning with genes, groups, and machines. At the descriptive level of analysis, they replace the term response with that of *state transition*. In the example above with the rat, a change from exploring the cage to pressing the lever could be described as such a state transition (moving from a state of exploration to a state of lever pressing). Importantly, states, as defined from this perspective, are used to describe state transitions. A behavior is then defined as a state transition in relation to one or more stimuli, for example, in relation to the music being played. This is once again a functional definition - the behavior is a function of stimuli. Learning is still defined as changes in behavior (state transitions in relation to stimuli) that occurs due to regularities (De Houwer and Hughes, 2022).

Based on this conceptual change a *system* is defined as a construct from the perspective of an observer, as sets of states that can be used to describe change. In common language we often refer to a rat "being a system", but technically it is rather that the system is a collection of topographical descriptions of a rat's physiological responses. With computer systems they could typically be described as collections of interdependent systems (Hayes and Fryling, 2018). For example, a robot taking part of an experimental task, might, for example, be described as hardware movements (like the robot arm) that are dependent on sensory equipment and on software interpreting those sensory inputs.

2.5 Human-level intelligence from an operant perspective

All above examples involve learning that are animal-level in the sense that they could be observed with an animal like a rat. The fact that operant conditioning can be observed with both humans and animals does not, however, mean that these processes are irrelevant for achieving human-level intelligence with artificial systems. On the contrary, we would argue that mechanisms enabling operant conditioning at the core (as with OpenNARS for Applications; Hammer, 2022) could very well be integral to the development of complex cognitive behaviors with AGI systems.

However, there might be limitations to the purely operant account as presented by Skinner. Already in the 1950s, critiques of Skinner's theories emerged, particularly in the context of language acquisition, arguing that verbal behavior requires more than just operant conditioning (Chomsky, 1959). Given this, the development of human-level AI from the perspective of operant conditioning could risk to be limited when it comes to higher-order functions such as language. These critiques have been thoroughly addressed by contemporary functional learning

theories of language and cognition, such as Relational Frame Theory (RFT; Hayes et al., 2001).

RFT posits that the foundation of human language and cognition lies in the ability to relate objects and events in arbitrary ways, an ability referred to as Arbitrarily Applicable Relational Responding (AARR; Hayes et al., 2001; Johansson, 2019). AARR allows individuals to respond to one stimulus in terms of another based on arbitrary contextual cues, rather than solely on the physical properties of the stimuli themselves. For example, learning to relate to an object and a spoken word as equivalent, despite having no inherent physical similarity, showcases this ability. Or, responding to a small coin as more valuable than a larger one, is another demonstration of AARR.

Importantly, such patterns of AARR (called "relational frames" in RFT) are assumed to be operant behaviors in themselves, which are learned through interaction with the environment and are subject to reinforcement (Hayes et al., 2021). Learning to derive relations in accordance with, for example, similarity, opposition, comparision, etc., seems to enable the development of complex cognitive skills such as language understanding, problem-solving, and abstract reasoning. Hence, from the perspective of RFT, intelligence as a whole can be viewed as a collection of interrelated relational frames that are dynamically shaped and modified through continuous engagement with one's environment (Cassidy et al., 2016; Hayes et al., 2021).

The implications of RFT for AGI research are profound. Since AARR is assumed to be learned behavior, it suggests that AGI systems could potentially achieve human-like intelligence through extensive training on relational tasks. There is a large amount of experimental RFT studies regarding training of relational framing abilities with humans across a variety of contexts and populations, providing a rich dataset to inform AGI development (Dixon et al., 2014; Cassidy et al., 2016). Also, an implication of an RFT perspective is the necessity for AGI systems to have mechanisms to learn relational frames from interactions with the environment in a manner similar to how humans learn throughout their lifetime. This means that AGI systems must be designed with architectures capable of not just operant conditioning in a traditional sense, but also with the ability to derive and apply relational frames dynamically.

A roadmap to AGI from the perspective of functional learning psychology and RFT would clearly emphasize operant conditioning abilities at the core, as suggested in this paper. Furthermore, it would advocate for trainings of increasingly complex AARR, in order to foster the development of advanced cognitive abilities (Johansson, 2020). Hence, such a roadmap could provide a clear specification of the requisite stages and milestones necessary for the development of AGI, aligning with the principles of Relational Frame Theory (RFT) (Hayes et al., 2001; 2021).

3 Non-axiomatic reasoning systems

A Non-Axiomatic Reasoning System (NARS) is a type of artificial intelligence system that operates under the assumption of insufficient knowledge and resources (AIKR) (Wang, 1995; 2006; 2013). The AIKR principle dictates that the system must function effectively despite having limited information and computational resources, a scenario that closely mirrors real-world conditions and human cognitive constraints.

All NARS systems implement a Non-Axiomatic Logic (NAL) (Wang, 2013), a term logic designed to handle uncertainty using experience-grounded truth values. Most NARS systems also makes use of concept-centric memory structure, which organizes the system's memory based on terms and subterms from the logic statements, leading to a more effective control of the inference process. Furthermore, all NARS systems use a formal language *Narsese*, that allows encoding of complex information and communication of NAL sentences within and between NARS systems.

3.1 Core principles of NARS

NARS systems are built on a few key concepts that distinguish them from traditional AI systems (Wang, 2022).

1. **Adaptation Under AIKR:** Unlike systems that assume abundant knowledge and resources, NARS thrives under constraints. It manages finite processing power and storage, operates in real-time, and handles tasks with varying content and urgency. This adaptability ensures that NARS remains relevant in dynamic and unpredictable environments.
2. **Experience-Based Learning and Reasoning:** Central to NARS is its concept-centered representation of knowledge. Concepts in NARS are data structures with unique identifiers, linked through relations such as inheritance, similarity, implication, and equivalence. These relationships are context-sensitive and derived from the system's experiences, allowing NARS to continuously update and refine its knowledge base as it encounters new information.
3. **Non-Axiomatic Logic:** Traditional AI often relies on axiomatic systems where certain truths are taken as given. In contrast, NARS employs non-axiomatic logic, where all knowledge is subject to revision based on new experiences. This approach supports a variety of inference methods, including deduction, induction, abduction, and analogy, enabling NARS to reason in a manner that is both flexible and grounded in empirical evidence.

3.2 Problem-solving and learning

NARS processes three types of tasks: incorporating new knowledge, achieving goals, and answering questions. It uses both forward and backward reasoning to handle these tasks, dynamically allocating its limited resources based on task priorities. This approach, known as case-by-case problem-solving, means that NARS does not rely on predefined algorithms for specific problems. Instead, it adapts to the situation at hand, providing solutions that are contextually appropriate and continuously refined.

Learning in NARS is a self-organizing process (Wang, 2022). The system builds and adjusts its memory structure—a network of interconnected concepts—based on its experiences. This structure evolves over time, allowing NARS to integrate new knowledge, resolve conflicts, and improve its problem-solving capabilities.

Unlike many machine learning models that require large datasets and extensive training, NARS learns incrementally in interaction with its environment and can accept inputs at various levels of abstraction, from raw sensorimotor data to complex linguistic information.

3.3 A unified cognitive model

One of the most significant aspects of NARS is its unified approach to cognitive functions. In NARS, reasoning, learning, planning, and perception are not separate processes but different manifestations of the same underlying mechanism. This integration provides a coherent framework for understanding and developing general intelligence, making NARS a versatile tool for a wide range of AI applications (Wang, 2022).

In conclusion, NARS represents a significant departure from conventional AI paradigms by embracing the challenges of limited knowledge and resources. Its unique combination of non-axiomatic reasoning, experience-based learning, and adaptive problem-solving positions NARS as a robust model for advancing artificial general intelligence.

4 OpenNARS for applications

OpenNARS for Applications (ONA) is a highly effective implementation of a NARS, designed to be suitable for practical applications such as robotics (Hammer and Lofthouse, 2020). At the core of ONA lies sensorimotor reasoning, which integrates sensory processing with motor actions to enable goal-directed behavior under conditions of uncertainty and limited resources. ONA differs from other NARS systems in several key aspects, including.

1. **Event-Driven Control Process:** ONA incorporates an event-driven control mechanism that departs from the more probabilistic and bag-based approach used in traditional NARS implementations, such as OpenNARS (Lofthouse, 2019). This shift allows ONA to prioritize processing based on the immediacy and relevance of incoming data and tasks. The event-driven approach is particularly advantageous in dynamic environments where responses to changes must be timely and context-sensitive.
2. **Separation of Sensorimotor and Semantic Inference:** Unlike other NARS models that often blend various reasoning functions, ONA distinctly separates sensorimotor inference from semantic inference (Hammer and Lofthouse, 2020). This division allows for specialized handling of different types of reasoning tasks—sensorimotor inference can manage real-time, action-oriented processes, while semantic inference deals with abstract, knowledge-based reasoning. This separation helps to optimize processing efficiency and reduces the computational complexity involved in handling diverse reasoning tasks simultaneously.
3. **Resource Management:** ONA places a strong emphasis on managing computational resources effectively, adhering to the Assumption of Insufficient Knowledge and Resources (AIKR). It is designed to operate within strict memory and processing constraints, employing mechanisms like priority-based forgetting and constant-time inference cycles. These features ensure that ONA can function continuously in resource-limited settings by efficiently managing its cognitive load and memory usage.
4. **Advanced Data Structures and Memory Management:** ONA utilizes a sophisticated system of data structures that include events, concepts, implications, and a priority queue system for managing these elements. This setup facilitates more refined control over memory and processing, prioritizing elements that are most relevant to the system's current goals and tasks. It also helps in maintaining the system's performance by managing the complexity and volume of information it handles.
5. **Practical Application Focus:** The architectural and control changes in ONA are driven by a focus on practical application needs, which demand reliability and adaptability. ONA is tailored to function effectively in real-world settings that require autonomous decision-making and adaptation to changing environments, making it more applicable and robust than its predecessors for tasks in complex, dynamic scenarios Hammer and Lofthouse (2020).

4.1 The architecture of ONA

ONA's architecture is composed of several interrelated components that work together to process sensory input, manage knowledge, make decisions, and learn from experience. These components are designed to handle the dynamic and uncertain nature of real-world environments, ensuring that the system can adapt and respond effectively (Hammer, 2022). The architecture is illustrated in Figure 1. ONA's architecture has a number of key components: 1) Event Providers, 2) FIFO Sequencer, 3) Cycling Events Queue, 4) Concept Memory, 5) Sensorimotor Inference Block, and 6) Declarative Inference Block. Each of these components plays a crucial role in ONA's operation, as described in detail below.

4.1.1 Event providers

Event providers are responsible for processing sensory inputs from various modalities, converting raw data into structured statements that the reasoning system can interpret. Each event provider is specialized for different types of sensory information, such as visual, auditory, or tactile data. These providers ensure that all relevant environmental information is captured and encoded as events, which are then fed into the system for further processing. The main functionality can be summarized as follows.

- **Sensor Data Processing:** Event providers preprocess raw sensor data to filter noise and extract meaningful information.
- **Event Encoding:** The processed data is encoded into statements or events that can be understood by the ONA system.

4.1.2 FIFO sequencer

The FIFO (First-In-First-Out) Sequencer maintains a sliding window of recent events. This component is essential for building and strengthening temporal implication links, which are used to

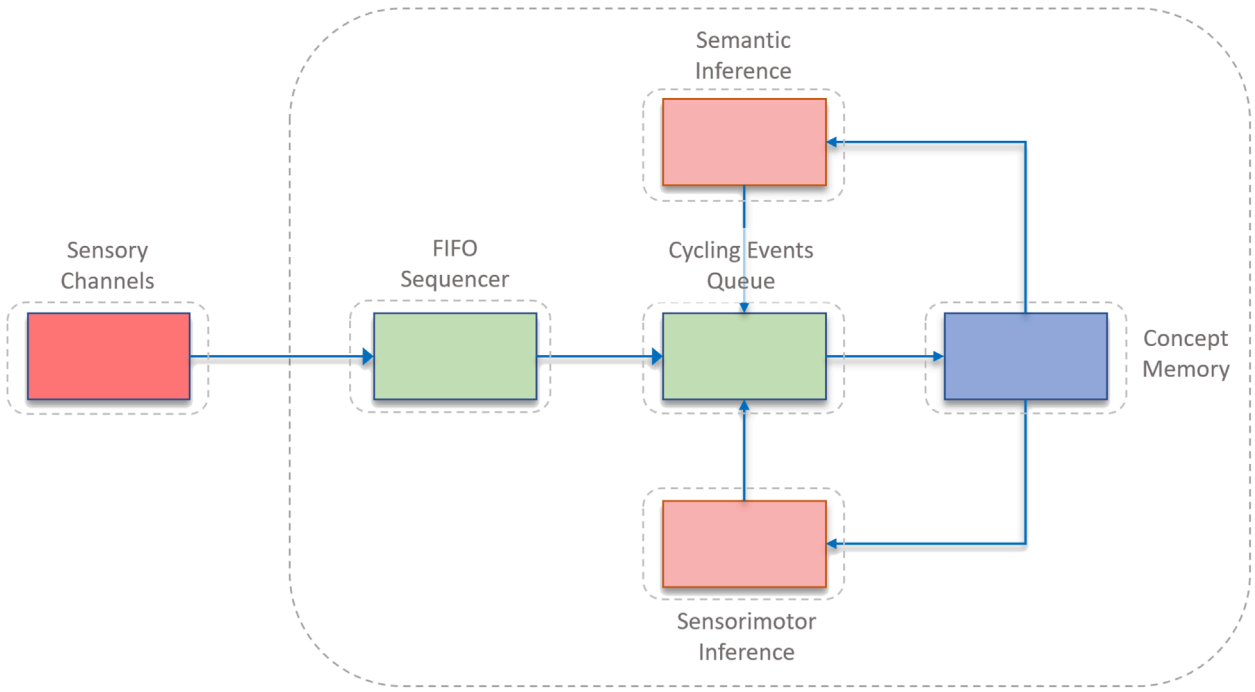


FIGURE 1
An overview of the architecture in OpenNARS for Applications (ONA). Reprinted with permission from Patrick Hammer, the author of ONA.

understand the sequence of events and their relationships over time. By keeping track of the recent history, the FIFO Sequencer allows ONA to form hypotheses about temporal patterns and causal relationships. As a note, in recent versions of ONA, the FIFO was removed and replaced by an explicit temporal inference block. This design is however not yet described in any scientific publications, and therefore the design with the FIFO has been described. The main functionality can be summarized as follows.

- **Event Sequencing:** It organizes events in a chronological order, maintaining a window of the most recent events.
- **Temporal Implications:** Builds and strengthens links between events based on their temporal proximity and sequence.

4.1.3 Cycling Events Queue

The Cycling Events Queue is a priority queue that serves as the central attention buffer of the system. All input and derived statements enter this queue, but only a subset can be selected for processing within a given timeframe due to the fixed capacity of the queue. This mechanism ensures that the most relevant and urgent information is processed first, while less critical information is discarded or delayed. The main functionality can be summarized as follows.

- **Priority Management:** Events are prioritized based on their importance and relevance to current goals.
- **Attention Focus:** Ensures that the system's limited processing resources are focused on the most critical tasks.

4.1.4 Concept Memory

Concept Memory acts as the long-term memory of the ONA system. It stores temporal hypotheses and supports their strengthening or weakening based on prediction success. This memory component allows ONA to retain knowledge over long periods, enabling cumulative learning and the ability to recall past experiences to inform current decision-making. The main functionality can be summarized as follows.

- **Hypothesis Management:** Stores and manages temporal and procedural hypotheses about the environment.
- **Evidence Accumulation:** Strengthens or weakens stored hypotheses based on new evidence and prediction outcomes.

4.1.5 Sensorimotor Inference Block

The Sensorimotor Inference Block is responsible for handling decision-making and subgoaling processes for goal events selected from the Cycling Events Queue. This component invokes algorithms for goal achievement, generating actions or subgoals that guide the system's behavior towards fulfilling its objectives. The main functionality can be summarized as follows.

- **Decision Making:** Selects the best actions to achieve current goals based on stored knowledge and recent events.
- **Subgoaling:** Decomposes complex goals into manageable subgoals, facilitating step-by-step achievement of objectives.

4.1.6 Declarative Inference Block

The Declarative Inference Block is responsible for higher-level reasoning tasks such as feature association, prototype formation, and relational reasoning. It utilizes human-provided knowledge to enhance the system's understanding of the environment and improve its reasoning capabilities. Though not utilized in the specific experiments described in the paper, this block is crucial for applications requiring complex knowledge integration and abstract reasoning. The main functionality can be summarized as follows.

- **Feature Association:** Links related features and concepts based on observed patterns and external knowledge.
- **Relational Reasoning:** Understands and reasons about relationships between different concepts and entities.

4.2 The operations of ONA

The main operations of ONA will be described below.

4.2.1 Truth value calculation

Truth values in ONA are based on positive and negative evidence supporting or refuting a statement, respectively. The system uses two measures: frequency (the ratio of positive evidence to total evidence) and confidence (the ratio of total evidence to total evidence plus one). This approach allows ONA to represent degrees of belief, accommodating the inherent uncertainty in real-world information.

The calculation of frequency and confidence is conducted as follows. Frequency: $f = \frac{w^+}{w}$, and Confidence: $c = \frac{w}{w+1}$, where w is the total amount of evidence, and w^+ is the positive evidence.

These values are used to evaluate the truth of implications and guide decision-making processes, ensuring that actions are based on the most reliable and relevant information available.

4.2.2 Implications and learning

ONA forms temporal and procedural implications through induction and revises them based on new evidence. Temporal implications represent sequences of events, while procedural implications represent action-outcome relationships. Learning involves accumulating positive and negative evidence for these implications and adjusting their truth values accordingly. Also if an implication exists (for example, `<(<A1 –> [left]> &/ ^left) =/> G>`, and $A1$ was observed followed by `^left`, an assumption of failure will be applied to the implication for implicit anticipation. This means, if the anticipation fails, the truth of the implication will be reduced by the addition of negative evidence, via an implicit negative G event, while the truth will increase due to positive evidence in case G happened.

The learning process at the core consists of.

- **Event Sequences:** Implications are formed when related events occur within the sliding window maintained by the FIFO Sequencer.
- **Evidence Update:** Positive and negative evidence is accumulated and used to revise the truth values of implications, ensuring they reflect the system's experiential knowledge.

4.2.3 Decision making and subgoaling

ONA's decision-making process is goal-driven, leveraging its knowledge of temporal and procedural implications to select actions or generate subgoals. The system evaluates the desire value of goals and subgoals, prioritizing them based on their likelihood of success and relevance to current objectives. The decision process can be described as follows.

- **Goal Deduction and Evaluation:** Determines the most desirable actions or subgoals based on stored implications and recent events.
- **Subgoal Generation:** Breaks down complex goals into smaller, manageable subgoals, facilitating efficient achievement through step-by-step actions.

4.2.4 Motor babbling

To trigger executions when no procedural knowledge yet exists, ONA periodically invokes random operations, a process called *Motor Babbling*. This enables ONA to execute operations despite any procedural knowledge that applies. Without this ability, ONA would not be able to do its initial steps of learning procedural knowledge (Hammer and Lofthouse, 2020).

4.3 Conclusion

The architecture of ONA integrates various components that collectively enable it to reason, learn, and make decisions under conditions of uncertainty and resource constraints. By demonstrating aspects of natural intelligence, such as continuous learning and goal-driven behavior, ONA offers a robust framework for developing intelligent systems capable of adapting to the complexities of real-world environments.

5 Related work

While we are not aware of any other attempt to integrate functional learning psychology with the Non-Axiomatic Reasoning System (NARS), there are several approaches that aim to implement the "biological basis" of operant conditioning using computational modeling or similar approaches. Importantly, it seems like most of these attempts take a mechanistic approach to operant conditioning, rather than a functional approach (as in this paper). Reinforcement learning, particularly model-free methods like Q-Learning and Deep Q-Networks (DQN), has gained significant attention for its ability to learn optimal policies through interactions with the environment (Mnih et al., 2015). These methods rely on the Markov property, where the next state depends only on the current state and action, simplifying the learning process but also limiting the system's ability to handle non-Markovian environments.

ONA diverges from RL by adopting a reasoning-based approach grounded in Non-Axiomatic Logic (NAL). Unlike RL, which optimizes a predefined reward function, ONA emphasizes real-time reasoning under uncertainty, adapting to insufficient knowledge and resources (Wang, 2013). This allows ONA to handle complex, non-Markovian environments more effectively. While RL methods struggle with sparse rewards and require extensive data to learn,

ONA leverages its reasoning capabilities to infer causal relationships and plan actions based on partial knowledge, making it more data-efficient (Hammer, 2022).

In summary, while reinforcement learning remains a powerful tool for specific, well-defined tasks, ONA offers a robust alternative for more complex, real-time applications. Its integration of reasoning under uncertainty, goal-driven learning, and adaptability positions it as a significant advancement in the quest for generalizable and resilient AI systems (Hammer, 2022).

6 Machine Psychology

Machine Psychology is an interdisciplinary framework for advancing AGI research. It aims to integrate principles from operant learning psychology (as described in Section 2), with the theory and implementation of NARS (as described in Sections 3 and 4). At the core of the integration is the assumption that adaptation is fundamental to intelligence, both biological and artificial.

Generally, Machine Psychology can be said to be a *functional* approach (as defined in Section 2 to the problem of building an AGI system. With this, we mean that the Machine Psychology framework enables the possibility to not only study functional relations between changes in the environment and changes in behavior (as in operant psychology), but also to study functional relations between mechanisms and changes in behavior. Hence, both experience of the system, and its mechanisms could in principle be manipulated.

In the case with studying operant conditioning with NARS, it means that it is indeed possible to both manipulate the system's experience, but also, in principle, to manipulate the mechanisms that are available (or not) during an experimental task.

This interdisciplinary approach might be likened to Psychobiology, that integrates psychology and biology (Dewsbury, 1991). Psychobiology is an interdisciplinary field that integrates biological and psychological perspectives to study the dynamic processes governing behavior and mental functions in whole, integrated organisms. It emphasizes the interaction between biological systems, such as the nervous and endocrine systems, and psychological phenomena, such as cognition, emotion, and behavior. This approach allows for the dual manipulation of factors related to both experience (as in psychology) and biological processes (as in biology) within a unified framework. By doing so, psychobiology provides a comprehensive understanding of how environmental and experiential factors can influence biological states and how biological conditions can shape psychological experiences, thus bridging the gap between the two domains to offer holistic insights into human and animal behavior (Dewsbury, 1991; Ritz and von Leupoldt, 2023).

Hence, one way to describe Machine Psychology, is that it is to computer science (and particularly NARS theory), as what Psychobiology is to biology.

6.1 An interaction with NARS

Within a Machine Psychology approach to NARS, it is possible to interact with NARS as one would do with an organism in psychological research in general. This will be illustrated in this section.

An example interaction can be described as follows, where each line ends with `:|:`, indicating temporal statements. First, a nonsense symbol *A*1 is presented to the left. Then, *A*2 is presented to the right. After that, the event *G* is established as something desirable by the system (the `!` indicates that the system desires the event). This triggers the system to execute an operation (for example, the operation `^left`). The researcher could then provide the event *G* as a consequence, leading to a derived contingency statement by the system. The entire interaction can be described as follows (where `//` represents comments):

```
   <A1 --> [left]>. :|: // A1 is presented
to the left
   <A2 --> [right]>. :|: // A2 is presented
to the right
   G! :|: // G is established as a desired
event
   ^left. :|: // ^left executed by the system
   G. :|: // G is provided as a consequence
   <(<A1 --> [left]> &/ ^left) =/> G> //
Derived by the system
```

This example aims to provide an example of how the researcher might interact with NARS, as if it was a biological organism. The researchers presents events, and the system responds, and the researcher once again presents an event as a consequence.

As part of this study, all interactions with ONA was done via its Python interface. Specifically, experimental designs was conducted in the Python-based open source experimental software OpenSesame, that was configured to interact with ONA (Mathôt et al., 2012; Mathôt and March 2022).

6.2 Learning psychology with NARS

Given the above example, it should be clear that interactions between NARS and its environment can be analyzed using the terms provided by functional learning psychology, as provided in Section 2. The operant learning examples are described at the descriptive level (the level of procedure), and learning effects can be described at the functional level. The analog to the cognitive/mechanistic level is the operations of the NARS system, as, for example, described in Sedtion 4.2. The three-term contingency (as described in Section 2.1) can be used to describe relations of events, operations and consequences. In the example above, the event `<A1 –> [left]>` functions as a discriminative stimulus, `^left` is a response, and `G` functions as a reinforcer. Importantly though, the `G!` (that establishes G as desired event) functions as an establishing operation.

7 Methods

7.1 OpenNARS for applications

The study used a version of OpenNARS for Applications (ONA) compiled with the parameter `SEMANTIC_INFERENCE_NAL_LEVEL` set to `0`, which means that only sensorimotor reasoning

were to be used. Hence, no declarative inference rules were available during the experiments.

For all three experiments, ONA was configured at starting time in the following way:

```
*babblingops=2
*motorbabbling=0.9
*setopname 1 ^left
*setopname 2 ^right
*volume=100
```

This indicates that ONA was set to have two operators `^left` and `^right`, and an initial chance of 90% for motor babbling.

7.2 Encoding of experimental setup

All experimental tasks were presented as temporal Narsese statements, as indicated by the `:|:` markers below. An arbitrary goal event `G! :|:` was presented at the end to trigger the execution of one of the two procedural operations `^left` and `^right` (through motor babbling or a decision). During training, feedback was given in the form of `G. :|:` (meaning to reinforce a correct choice) or `G. :|: {0.0 0.9}` (to indicate that the system had conducted an incorrect choice). Between each trial, 100 time steps was entered, by feeding `100` to ONA.

```
<A1 --> [sample]>. :|:
<B1 --> [left]>. :|:
<B2 --> [right]>. :|:
G! :|:
```

The first three lines are so-called inheritance statements, with properties on the right-hand side, indicating that the events $A1$, $B1$ and $B2$ are either on the left, right or at the position of a sample.

7.3 Experimental designs

In this section, the experimental designs will be detailed of the three tasks: 1) The simple discrimination task, 2) The changing contingencies task, and 3) The conditional discriminations task. The tasks are further illustrated with a few examples in Figure 2.

The first experiment investigated if NARS could learn in the form of operant conditioning, specifically in the form of simple discriminations. In the experiment, three phases were used: Baseline assessment, Training (with feedback), and Testing (without feedback). In all phases, training and testing were done in blocks of trials. One trial could, for example, be that $A1$ was to the left, and $A2$ was to the right. A block contained twelve trials, with the two possible trials possible (depending on the location of $A1$ and $A2$), each presented six times in random order.

1. **Baseline:** During the baseline assessment, which was three blocks, no feedback was given. This phase was included to establish a baseline probability of responding correct. It was expected that the system would respond correctly by chance in 50% of the trials.
2. **Training:** Then, the system was trained on a set of three blocks. Feedback was given when the system was correct (for example, executing `^left` when $A1$ was to the left), and when not correct.
3. **Testing:** The system was then tested (without feedback) on three blocks, with the contingencies that previously had been trained.

The second experiment investigated if ONA could adapt to changing conditions midway through the task. Five phases were used: Baseline, Training 1 (with feedback), Testing 1 (without feedback), Training 2 (with feedback), and Testing 2 (without feedback). All blocks contained twelve trials.

1. **Baseline:** Two blocks, where no feedback was given.
2. **Training 1:** Four blocks of, where feedback was given. This phase aimed to train the system in executing `^left` when $A1$ was to the left, and `^right` when $A1$ was to the right.
3. **Testing 1:** Then, the system was tested over two blocks (without feedback) on what was trained the previous phase.
4. **Training 2:** This phase of four blocks aimed to train in reversed contingencies compared to the first training. That is, the phase aimed to train ONA into executing `^left` when $A2$ was to the left (and hence $A1$ to the right), and `^right` when $A2$ was to the right.
5. **Testing 2:** Over two blocks, the system was tested, without feedback, on the contingencies trained in the previous phase.

Finally, in the third experiment, that investigated if the system could learn conditional discriminations, three phases were used: Baseline, Training, and Testing.

1. **Baseline:** Three blocks of 12 trials, where no feedback was given.
2. **Training:** Six blocks, where feedback was given. For example, when $A1$ was the sample, and $B1$ to the left, the system was reinforced for executing `^left`.
3. **Testing:** The system was then tested, without feedback, on three blocks of 12 trials, with the contingencies that previously had been trained.

8 Results

8.1 Simple discrimination task

During baseline, the amount of correct trials ranged between 0% and 50% during the three blocks, indicating that no learning happened. In the training phase, NARS was 100% correct on all trials already in the second out of three blocks, indicating a rapid learning. Finally, in the testing, where no feedback was provided, NARS performed consistently 100% correct across all three blocks of trials. The results are illustrated in Figure 3.

The average confidence values for the two target hypotheses went from 0.56 to 0.82. These two hypotheses were

```
<(<A1 --> [left]> &/ ^left) =/> G>
```

and

```
<(<A1 --> [right]> &/ ^right) =/> G>
```

The increase in average confidence value is also illustrated in Figure 3.

In summary, the results indicate that ONA indeed can learn in the form of operant conditioning.

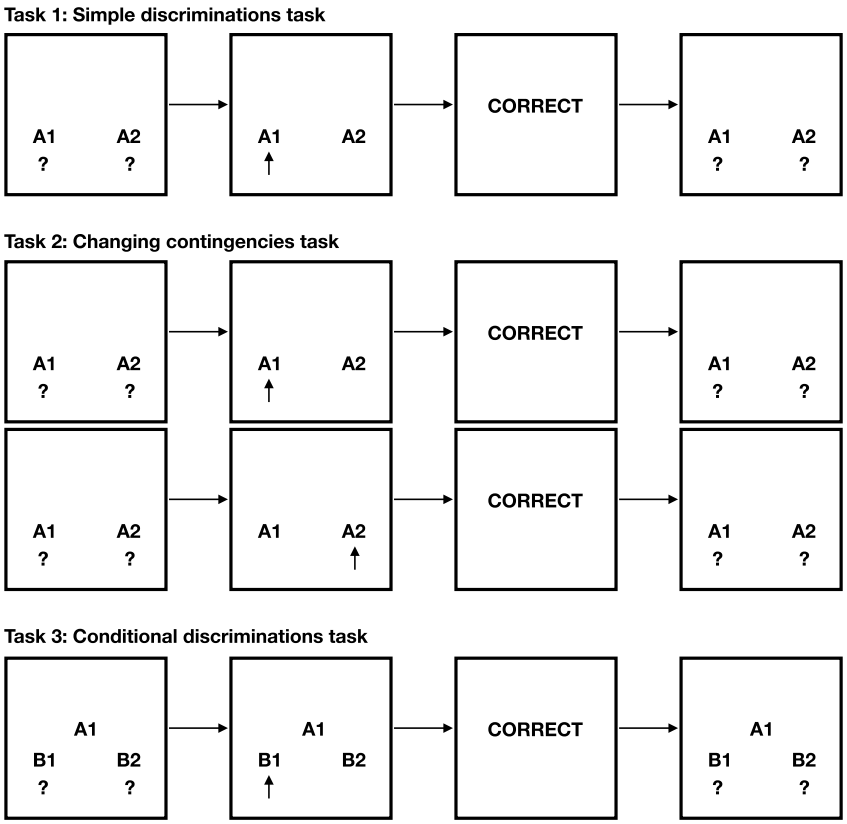


FIGURE 2
Examples from the three experimental tasks investigated.

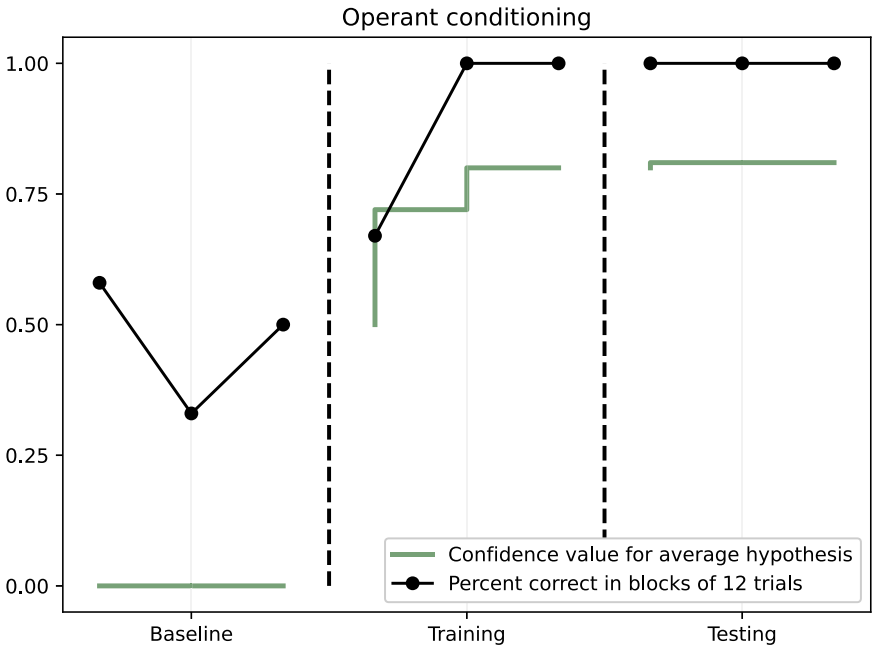


FIGURE 3
Operant conditioning. Dots illustrate the percent of correct in blocks of 12 trials. The solid line shows the mean NARS confidence value for hypotheses.

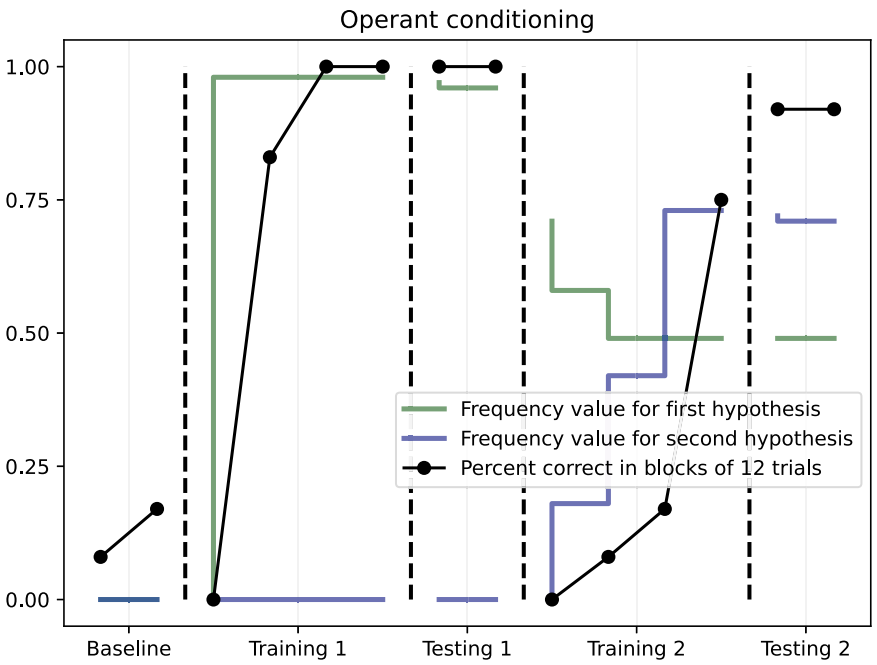


FIGURE 4
Operant conditioning with changing contingencies. Dots illustrate the percent of correct in blocks of 12 trials. The solid lines show the mean NARS frequency values for the respective hypotheses.

8.1.1 NARS examples from the training phase

A few example trials from the training session follows. Let's say that the system was exposed to the following NARS statements:

```
<A2 --> [left]>. :|:
<A1 --> [right]>. :|:
G! :|:
```

If it is early in the training, NARS might use Motor Babbling to execute the `^right` operation. Since this is considered correct in the experiment, the feedback `G. :|:` would be given to NARS, followed by 100 time steps. Only from this single interaction, NARS would form a hypothesis using *Temporal Induction*:

```
<(<A1 --> [right]> &/ ^right) =/> G>
```

When the same situation happens again later during the training phase, ONA will not rely on motor babbling, but instead use its decision making algorithm and *Goal Deduction*, as detailed by (Hammer, 2022).

8.2 Changing contingencies task

As expected, no learning happened during the baseline phase, where NARS was less than 25% correct in both phases. In the first training phase, NARS was 100% correct after two completed blocks of 12 trials. During testing, the system was 100% correct without any feedback being present. In the second training phase, where the contingencies were reversed, the system could adapt to the change as indicated by the increase in number of correct responses over time, with 75% correct in the final block of the phase. Finally, in the second testing phase, the system's performance was 91.7% correct, indicating that a successful retraining had been conducted. The results are further illustrated in Figure 4.

To further illustrate how the NARS system was able to adapt to changing contingencies, the change in average frequency value of the two target hypotheses can be described over time. This is also illustrated in Figure 4. As seen in Figure 4, the average frequency value for the first hypothesis was close to 1.0 during the first training and testing, meaning that the system had not received any negative evidence. However, when the contingencies were reversed in the second training phase, the frequency value of the first hypothesis immediately decreased, taking the negative evidence into account. The frequency value of the second hypothesis however, did not rise above zero until the start of the second training, where the hypothesis got positive evidence for the first time.

These results do all together indicate that a NARS system can adapt in realtime in the form that is necessary when contingencies are reversed midway through a task.

8.2.1 Examples from changed contingencies

The experiment starts out similar as to the example in Section 8.1.1 during the first training phase. However, after the contingencies change, and reinforcement is not provided for executing `^left` and `^right` when *A*1 is to the left and right, respectively, the system is forced to readapt. For example, if the following situation is shown to the system:

```
<A1 --> [left]>. :|:
<A2 --> [right]>. :|:
G! :|:
```

The system will execute `^left` based on its previous learning. However, instead of `G. :|:` as a consequence, `G. :|: {0.0 0.9}` will be provided. An explanation of how *Revision* is used will be provided.

Before the negative feedback, the following hypothesis will have a frequency close to 1.0:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.98, 0.41}
```

In the above, `0.98, 0.41` means frequency = 0.98, and confidence = 0.41.

However, with the negative feedback shown above, the following hypothesis will be derived:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.00, 0.19}
```

Together, these two hypothesis with different truth values will be revised as follows:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.74, 0.48}
```

When NARS combines the positive and negative evidence, the frequency value goes down from 0.98 to 0.74, and the confidence value goes up from 0.41 to 0.48, as the system has gained even more evidence and is more confident in its conclusions.

With repeated examples smilar to the above, the system will eventually go back to motor babbling, and `^right` will be executed, leading to a reinforcing consequence. That will lead to the following hypothesis being formed:

```
<(<A2 --> [right]> &/ ^right) =/> G>.
```

In summary, the mechanism of *Revision*, in combination to what have been covered previously, enables the system to adapt to changing contingencies.

8.3 Conditional discriminations task

As with the previous experiments, no learning happened during the three-block baseline. During training, NARS was more than 75% correct after two completed blocks of 12 trials. In the testing, NARS performed 100% correct, without feedback, across three blocks of trials. These results are illustrated in Figure 5.

The four target hypotheses were the following:

```
<((<A1 --> [sample]> &/ <B1 --> [left]>)
&/ ^left) =/> G>
<((<A1 --> [sample]> &/ <B1 --> [right]>)
&/ ^right) =/> G>
<((<A2 --> [sample]> &/ <B2 --> [left]>)
&/ ^left) =/> G>
<((<A2 --> [sample]> &/ <B2 --> [right]>)
&/ ^right) =/> G>
```

The average confidence value for these hypotheses increased from 0.13 to 0.70 during the training phase, as also illustrated in Figure 5.

8.3.1 NARS examples from conditional discrimination training

A few example trials from the training session follows. Let's say that the system was exposed to the following NARS statements:

```
<A1 --> [sample]>. :|:
<B2 --> [left]>. :|:
<B1 --> [right]>. :|:
G! :|:
```

If it is early in the training, NARS might use motor babbling to execute the `^right` operation. Since this is considered correct in the experiment, the feedback `G. :|:` would be given to NARS, followed by 100 time steps. From this single interaction, NARS would form a hypothesis:

```
<((<A1 --> [sample]> &/ <B1 --> [right]>)
&/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.15
```

Importantly, after this single trial, NARS would also form simpler hypothesis such as:

```
<(<B1 --> [right]> &/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.21
<(<A1 --> [sample]> &/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.16
```

This means, that if the same trial was to be presented again (all four possible trials will be presented three times in a block of twelve trials), NARS would respond `^right` again, but the decision being based on the simpler hypothesis, since that hypothesis has the highest confidence value.

Let's say, that within the same block of 12 trials, the next trial to be presented to NARS was the following:

```
<A1 --> [sample]>. :|:
<B1 --> [left]>. :|:
<B2 --> [right]>. :|:
G! :|:
```

NARS would initially respond `^right`, with the decision being made from the simple hypothesis `<(<A1 –> [sample]> &/ ^right) =/> G>.`

This would be considered wrong in the experiment, and the feedback `G. :|:{0.0 0.9}` would be given to NARS. This would lead to negative evidence for the simple hypothesis. If the same trial was presented again, NARS would then likely resort to motor babbling that could execute the `^left` operation. Over repeated trials with feedback, the simpler hypotheses would get more negative evidence, and the confidence values of the more complex target hypotheses would increase.

In summary, NARS can learn increasingly complex hypotheses, with repeated examples.

8.4 NARS mechanisms

Given the examples above, we will now provide further clarifications of the results in terms of mechanisms and inference rules that are implemented in ONA.

In all three tasks, the confidence increase followed from repeated examples which provide evidence to the respective target hypotheses. For this to happen and to derive the truth values, the following mechanisms in NARS were necessary.

1. **Temporal induction:** Given events that *A*1 is to the left, the `^left` operation, and *G*, then derive positive evidence for a relation like `<(<A1 –> [left]> &/ ^left) =/> G>`
2. **Goal deduction:** Given for example, `<(<A1 –> [left]> &/ ^left) =/> G>` and a precondition that *A*1 is to the left, and the event *G*!, then by deduction derive that the `^left` operation is to be executed.
3. **Motor babbling:** The ability to execute operations functions as the means for exploration in the sense that it enables the system to try out new things.

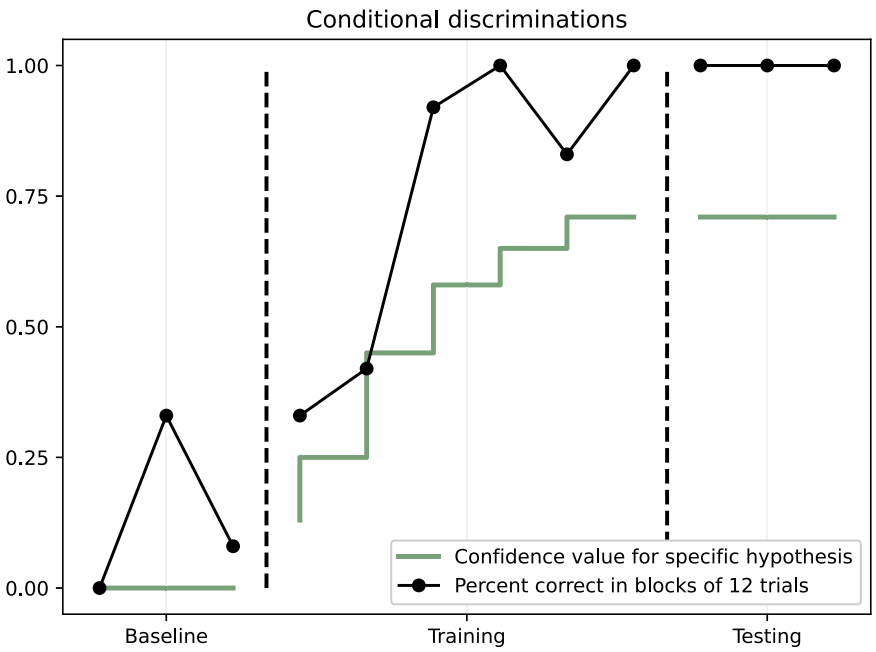


FIGURE 5
Conditional discriminations. Dots illustrate the percent of correct in blocks of 12 trials. The solid line shows the mean NARS confidence value for hypotheses.

4. **Anticipation:** To derive negative evidence to a hypothesis, based on that the antedecent happened but the consequent did not. For example, `<(<A1 –> [sample]> &/ ^right) =/> G>` can receive negative evidence based on anticipation.
5. **Revision:** To summarize the positive evidence and the negative evidence for a statement.

9 Discussion and conclusion

The results of this study demonstrate the feasibility and effectiveness of integrating principles from operant conditioning with the Non-Axiomatic Reasoning System (NARS) to advance the field of Artificial General Intelligence (AGI). This interdisciplinary framework, referred to as Machine Psychology, offers a novel approach to understanding and developing intelligent systems by emphasizing adaptation, a core aspect of both biological and artificial intelligence.

9.1 Summary of findings

The experiments conducted in this study aimed to evaluate the ability of NARS, specifically the OpenNARS for Applications (ONA) implementation, to perform operant conditioning tasks. The three tasks—simple discrimination, changing contingencies, and conditional discriminations—provided a comprehensive assessment of the system's learning and adaptation capabilities.

In the simple discrimination task, NARS demonstrated rapid learning, achieving 100% correct responses during the training phase and maintaining this performance in the testing phase without feedback. This indicates that NARS can effectively learn and adapt based on positive reinforcement, a key aspect of operant conditioning.

The changing contingencies task further highlighted the system's adaptability. When the contingencies were reversed midway through the task, NARS was able to adjust its behavior accordingly, showing a significant decrease in errors and an increase in correct responses during the retraining phase. This flexibility is crucial for AGI systems operating in dynamic environments where conditions can change unpredictably.

The conditional discriminations task showcased NARS's ability to handle more complex learning scenarios. Despite the increased difficulty, the system achieved high accuracy, indicating that it can form and utilize more intricate hypotheses based on conditional cues. This capability is essential for developing AGI systems that require sophisticated cognitive skills.

9.2 Implications for AGI research

The success of NARS in these operant conditioning tasks has several important implications for AGI research. First, it validates the use of learning psychology principles, particularly operant conditioning, as a guiding framework for developing intelligent systems. The results suggest that mechanisms enabling operant conditioning are integral to the development of adaptive behaviors and cognition in AGI systems.

Second, the experiments carried out as part of this study, can be said to constitute key milestones of AGI research, as has been suggested by us elsewhere (Johansson, 2020). Operant psychology research provides examples of increasingly complex tasks, that can

be used to test the abilities of an AGI system. The use of functional learning psychology principles to guide AGI research also enable metrics to be used to evaluate AGI systems, as demonstrated in this paper.

Third, the study highlights the potential of NARS as a robust model for AGI. Unlike traditional AI systems that rely on predefined algorithms and large datasets, NARS operates effectively under conditions of insufficient knowledge and resources. This adaptability makes it well-suited for real-world applications where information is often incomplete and environments are constantly changing.

Fourth, the integration of sensorimotor reasoning with operant conditioning principles in ONA provides a scalable and flexible framework for AGI development. By demonstrating aspects of natural intelligence, such as continuous learning and goal-driven behavior, ONA offers a practical approach to building intelligent systems that can interact with and learn from their environments in real-time.

9.3 Future directions

The findings of this study open several avenues for future research. One potential direction is to explore the integration of additional cognitive and behavioral principles from functional learning psychology into NARS. Future research can be guided by operant theories of cognition, such as Relational Frame Theory, as suggested by Johansson (2019).

Another important direction is to apply the Machine Psychology framework to more complex and diverse tasks beyond the idealized examples provided in this paper. By testing NARS in various real-world scenarios, such as autonomous robotics, natural language processing, and human-computer interaction, researchers can evaluate the system's generalizability and robustness across different domains.

Additionally, further refinement of the sensorimotor inference and declarative inference components in ONA could lead to improvements in the system's performance. Enhancing the efficiency of resource management, memory structures, and event-driven control processes will be critical for scaling up the system to handle more sophisticated tasks and larger datasets.

9.4 Conclusion

In conclusion, this study demonstrates that integrating operant conditioning principles with NARS offers a promising pathway for advancing AGI research. The Machine Psychology framework provides a coherent and experimentally grounded approach to studying and developing intelligent systems. By emphasizing adaptation and learning from environmental interactions, this interdisciplinary approach has the potential to significantly advance the field of AGI and bring us closer to achieving human-level artificial intelligence.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

RJ: Conceptualization, Formal Analysis, Investigation, Methodology, Writing–original draft, Writing–review and editing, Funding acquisition, Project administration, Visualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was in part financially supported by Digital Futures through grant agreement KTH-RPROJ-0146472.

Acknowledgments

The author would like to thank Patrick Hammer and Tony Lofthouse for many valuable discussions regarding the work presented in this paper.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision

Publisher's note

References

Brembs, B. (2003). Operant conditioning in invertebrates. *Curr. Opin. Neurobiol.* 13, 710–717. doi:10.1016/j.conb.2003.10.002

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., et al. (2023). *Sparks of artificial general intelligence: early experiments with gpt-4. arXiv preprint arXiv:2303.12712.*

Cassidy, S., Roche, B., Colbert, D., Stewart, I., and Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learn. Individ. Differ.* 47, 222–235. doi:10.1016/j.lindif.2016.03.001

Chiesa, M. (1994). *Radical behaviorism: the philosophy and the science.* (Authors Cooperative).

Chomsky, N. (1959). Verbal behavior. *Language* 35, 26–58. doi:10.2307/411334

De Houwer, J., Barnes-Holmes, D., and Moors, A. (2013). What is learning? on the nature and merits of a functional definition of learning. *Psychonomic Bull. & Rev.* 20, 631–642. doi:10.3758/s13423-013-0386-3

De Houwer, J., and Hughes, S. (2020). *The psychology of learning: an introduction from a functional-cognitive perspective*. MIT Press.

De Houwer, J., and Hughes, S. J. (2022). Learning in individual organisms, genes, machines, and groups: a new way of defining and relating learning in different systems. *Perspect. Psychol. Sci.* 18, 649–663. doi:10.1177/17456916221114886

Dewsbury, D. A. (1991). Psychobiology. *Am. Psychol.* 46, 198–205. doi:10.1037//0003-066x.46.3.198

Dixon, M. R., Whiting, S. W., Rowsey, K., and Belisly, J. (2014). Assessing the relationship between intelligence and the peak relational training system. *Res. Autism Spectr. Disord.* 8, 1208–1213. doi:10.1016/j.rasd.2014.05.005

Hammer, P. (2022). "Reasoning-learning systems based on non-axiomatic reasoning system theory," in *International workshop on self-supervised learning* (PMLR) 192, 89–107.

Hammer, P., Isaev, P., Lofthouse, T., and Johansson, R. (2023). "Ona for autonomous ros-based robots," in *Artificial general intelligence: 15th international conference, AGI 2022, Seattle, WA, USA, august 19–22, 2022, proceedings* (Springer), 231–242.

Hammer, P., and Lofthouse, T. (2020). "'opennars for applications': architecture and control," in *International conference on artificial general intelligence* (Springer), 193–204.

Hawkins, J. (2021). *A thousand brains: a new theory of intelligence*. (Hachette UK).

Hayes, L. J., and Fryling, M. J. (2018). Psychological events as integrated fields. *Psychol. Rec.* 68, 273–277. doi:10.1007/s40732-018-0274-3

Hayes, S. C., Barnes-Holmes, D., and Roche, B. (2001). *Relational frame theory: a post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.

Hayes, S. C., Law, S., Assemi, K., Falletta-Cowden, N., Shamblin, M., Burleigh, K., et al. (2021). Relating is an operant: a fly over of 35 years of rft research. *Perspect. em Análise do Comportamento* 12, 005–032. doi:10.18761/pac.2021.v12.rft.02

Hutter, M. (2004). *Universal artificial intelligence: sequential decisions based on algorithmic probability*. Springer Science & Business Media.

Johansson, R. (2019). "Arbitrarily applicable relational responding," in *International conference on artificial general intelligence* (Springer), 101–110.

Johansson, R. (2020). "Scientific progress in agi from the perspective of contemporary behavioral psychology," in *OpenNARS workshop, artificial general intelligence (AGI-2020)*.

Laird, J. E. (2019). *The Soar cognitive architecture*. MIT press.

Lashley, K. S. (1938). Conditional reactions in the rat. *J. Psychol.* 6, 311–324. doi:10.1080/00223980.1938.9917609

Lofthouse, T. (2019). "Alann: an event driven control mechanism for a non-axiomatic reasoning system (nars)," in *NARS2019 workshop at AGI*.

Mathôt, S., and March, J. (2022). Conducting linguistic experiments online with opensesame and osweb. *Lang. Learn.* 72, 1017–1048. doi:10.1111/lang.12509

Mathôt, S., Schreij, D., and Theeuwes, J. (2012). Opensesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. methods* 44, 314–324. doi:10.3758/s13428-011-0168-7

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *nature* 518, 529–533. doi:10.1038/nature14236

Ritz, T., and von Leupoldt, A. (2023). Introduction to the 2022 special issue on neuroscience and psychobiology of respiration in biological psychology. *Biol. Psychol.* 176, 108478. doi:10.1016/j.biopsycho.2022.108478

Roche, B., and Barnes, D. (1997). The behavior of organisms? *Psychol. Rec.* 47, 597–618. doi:10.1007/bf03395248

Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. New York, NY: Appleton-Century-Crofts.

Skinner, B. F. (1953). *Science and human behavior. 92904*. Simon & Schuster.

Stewart, I., and McElwee, J. (2009). Relational responding and conditional discrimination procedures: an apparent inconsistency and clarification. *Behav. Analyst* 32, 309–317. doi:10.1007/bf03392194

Wang, P. (1995). *Non-axiomatic reasoning system: exploring the essence of intelligence*. Indiana University.

Wang, P. (2006). *Rigid flexibility: the logic of intelligence*, 34. Springer Science & Business Media.

Wang, P. (2012). "Theories of artificial intelligence—meta-theoretical considerations," in *Theoretical foundations of artificial general intelligence* (Springer), 305–323.

Wang, P. (2013). *Non-axiomatic logic: a model of intelligent reasoning* (World Scientific).

Wang, P. (2019). On defining artificial intelligence. *J. Artif. General Intell.* 10, 1–37. doi:10.2478/jagi-2019-0002

Wang, P. (2022). "A Unified Model of Reasoning and Learning," in *Proceedings of the Second International Workshop on Self-Supervised Learning* (PMLR) 159, 28–48.

WMU Rat Lab(2009). *More stimulus discrimination*. Available at: https://youtu.be/S2SqKRCa1a8(Accessed October 3, 2014).

frontiers | Frontiers in Robotics and AI

TYPE Original Research
PUBLISHED 12 November 2024
DOI 10.3389/frobt.2024.1384609

OPEN ACCESS

EDITED BY
Tony J. Prescott,
The University of Sheffield, United Kingdom

REVIEWED BY
José Antonio Becerra Permuy,
University of A Coruña, Spain
Yara Khaluf,
Wageningen University and Research, Netherlands

*CORRESPONDENCE
Silvia Ferrari,
ferrari@cornell.edu

RECEIVED 09 February 2024
ACCEPTED 19 August 2024
PUBLISHED 12 November 2024

CITATION
Chen Y, Zhu P, Alers A, Egner T, Sommer MA and Ferrari S (2024) Heuristic satisficing inferential decision making in human and robot active perception.
Front. Robot. AI 11:1384609.
doi: 10.3389/frobt.2024.1384609

Heuristic satisficing inferential decision making in human and robot active perception

Yucheng Chen[1], Pingping Zhu[2], Anthony Alers[3], Tobias Egner[4], Marc A. Sommer[3] and Silvia Ferrari[1]*

[1]Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, NY, United States, [2]College of Engineering and Computer Sciences, Marshall University, Huntington, IN, United States, [3]Department of Biomedical Engineering (BME), Duke University, Durham, NC, United States, [4]Center fir Cognitive Neuroscience, Duke Institute for Brain Sciences, Duke University, Durham, NC, United States

Inferential decision-making algorithms typically assume that an underlying probabilistic model of decision alternatives and outcomes may be learned *a priori* or online. Furthermore, when applied to robots in real-world settings they often perform unsatisfactorily or fail to accomplish the necessary tasks because this assumption is violated and/or because they experience unanticipated external pressures and constraints. Cognitive studies presented in this and other papers show that humans cope with complex and unknown settings by modulating between near-optimal and satisficing solutions, including heuristics, by leveraging information value of available environmental cues that are possibly redundant. Using the benchmark inferential decision problem known as "treasure hunt", this paper develops a general approach for investigating and modeling active perception solutions under pressure. By simulating treasure hunt problems in virtual worlds, our approach learns generalizable strategies from high performers that, when applied to robots, allow them to modulate between optimal and heuristic solutions on the basis of external pressures and probabilistic models, if and when available. The result is a suite of active perception algorithms for camera-equipped robots that outperform treasure-hunt solutions obtained via cell decomposition, information roadmap, and information potential algorithms, in both high-fidelity numerical simulations and physical experiments. The effectiveness of the new active perception strategies is demonstrated under a broad range of unanticipated conditions that cause existing algorithms to fail to complete the search for treasures, such as unmodelled time constraints, resource constraints, and adverse weather (fog).

KEYWORDS

satisficing, heuristics, active perception, human, studies, decision-making, treasure hunt, sensor

1 Introduction

Rational inferential decision-making theories obtained from human or robot studies to date assume that a model me be used either off-line or on-line in order to compute satisficing strategies that maximize appropriate utility functions and/or satisfy given mathematical constraints (Simon, 1955; Herbert, 1979; Caplin and Glimcher, 2014; Nicolaides, 1988; Simon, 2019). When a probabilistic world model is available, for example, methods such as optimal control, cell decomposition, probabilistic roadmaps,

and maximum utility theories, may be applied to inferential decision-making problems such as robot active perception, planning, and feedback control (Fishburn, 1981; Lebedev et al., 2005; Scott, 2004; Todorov and Jordan, 2002; Ferrari and Wettergren, 2021; Latombe, 2012; LaValle, 2006). In particular, active perception, namely, the ability to plan and select behaviors that optimize the information extracted from the sensor data in a particular environment, has broad and extensible applications in robotics that also highlights human abilities to make decisions when only partial or imperfect information is available.

Many "model-free" reinforcement learning (RL) and approximate dynamic programming (ADP) approaches have also been developed on the basis of the assumption that a partial or imperfect model is available in order to predict the next system state and/or "cost-to-go", and optimize the immediate and potential future rewards, such as information value (Bertsekas, 2012; Si et al., 2004; Powell, 2007; Ferrari and Cai, 2009; Sutton and Barto, 2018; Wiering and Van Otterlo, 2012; Abdulsaheb and Kadhim, 2023). Given the computational burden carried by learning-based methods, various approximations have also been proposed. For instance, approximate dynamic programming (ADP) methods have been developed based on the assumption that a partial or imperfect model is available to predict the next system state and/or "cost-to-go." These methods aim to optimize immediate and potential future rewards, such as information value (Bertsekas, 2012; Si et al., 2004; Powell, 2007; Ferrari and Cai, 2009; Sutton and Barto, 2018; Wiering and Van Otterlo, 2012), typically also exploiting world models available *a priori* in order to predict the next world state.

Other machine learning (ML) and artificial intelligence (AI) methods can be broadly categorized into two fundamental learning-based approaches. The first approach is deep reinforcement learning (DRL), where models incorporate classical Markov decision process theories and use a human-crafted or data-extracted reward function to train an agent to maximize the probability of gaining the highest reward (Silver et al., 2014; Lillicrap et al., 2015; Schulman et al., 2017). The second approach follows the learning from demonstration paradigm, also known as imitation learning (Chen et al., 2020; Ho and Ermon, 2016). Because of their need for extensive and domain-specific data, data-driven methods are also not typically applicable to situations that cannot be foreseen *a priori*.

Given the ability of natural organisms to cope with uncertainty and adapt to unforeseen circumstances, a parallel thread of development has focused on biologically inspired models, especially for perception-based decision making. These methods are typically computationally highly efficient and include motivational models, which use psychological motivations as incentives for agent behaviors (Lewis and Cañamero, 2016; O'Brien and Arkin, 2020; Lones et al., 2014), cognitive models, which transfer human mental and emotional functions into robots (Vallverdú et al., 2016; Martin-Rico et al., 2020). The implementation of cognitive models are usually in the form of heuristics, and their applications range from energy level maintenance (Batta and Stephens, 2019) to domestic environment navigation (Kirsch, 2016).

Humans have also been shown to use internal world models for inferential decision-making whenever possible, a characteristic first referred to as "substantial rationality" in (Simon, 1955; Herbert, 1979). As also shown by the human studies on passive and active satisficing perception presented in this paper, given sufficient data, time, and informational resources, a globally rational human decision-maker uses an internal model of available alternatives, probabilities, and decision consequences to optimize both decision and information value in what is known as a "small-world" paradigm (Savage, 1972). In contrast, in "large-world" scenarios, decision-makers face environmental pressures that prevent them from building an internal model or quantifying rewards, because of pressures such as missing data, time and computational power constraints, or sensory deprivation, yet still manage to complete tasks by using "bounded rationality" (Simon, 1997). Under these circumstances, optimization-based methods may not only be infeasible, returning no solution, but also cause disasters resulting from failing to take action (Gigerenzer and Gaissmaier, 2011). Furthermore, Simon and other psychologists have shown that humans can overcome these limitations in real life via "satisficing decisions" that modulate between near-optimal strategies and the use of heuristics to gather new information and arrive at fast and "good-enough" solutions to complete relevant tasks.

To develop satisficing solutions for active robot perception, herein, we consider here the class of sensing problems known as treasure hunt (Ferrari and Cai, 2009; Cai and Ferrari, 2009; Zhang et al., 2009; Zhang et al., 2011). The mathematical model of the problem, comprised of geometric and Bayesian network descriptions demonstrated in (Ferrari and Wettergren, 2021; Cai and Ferrari, 2009), is used to develop a new experimental design approach that ensures humans and robots experience the same distribution of treasure hunts in any given class, including time, cost, and environmental pressures inducing satisficing strategies. This novel approach enables not only the readily comparison of the human-robot performance but also the generalization of the learned strategies to any treasure hunt problem and robotic platform. Hence, satisficing strategies are modeled using human decision data obtained from passive and active satisficing experiments, ranging from desktop to virtual reality human studies sampled from the treasure hunt model. Subsequently, the new strategies are demonstrated through both simulated and physical experiments involving robots under time and cost pressures, or subject to sensory deprivation (fog).

The treasure hunt problem under pressure, formulated in Section 2. and referred to as satisficing treasure hunt herein, is an extension of the robot treasure hunt presented in Cai and Ferrari (2009); Zhang et al. (2009), which introduces motion planning and inference in the search for Spanish treasures originally used in Simon and Kadane (1975) to investigate satisficing decisions in humans. Whereas the search for Spanish treasures amounts to searching a (static) decision tree with hidden variables, the robot treasure hunt involves a sensor-equipped robot searching for targets in an obstacle-populated workspace. As shown in Ferrari and Wettergren (2021) and references therein, the robot treasure hunt paradigm is useful in many mobile sensing applications involving multi-target detection and classification. In particular, the problem highlights the coupling of action decisions that change the physical state of the robot (or decision-maker) with test decisions that allow the robot to gather information from the targets via onboard sensors. In this paper, the satisficing treasure hunt is introduced to investigate and model human satisficing perception strategies under external pressures in passive and active tasks, first via desktop simulations

and then in the Duke immersive Virtual Environment (DiVE) (Zielinski et al., 2013), as shown in Supplementary Figure S1.

To date, substantial research has been devoted to solving treasure hunt problems for many robots/sensor types, in applications as diverse as demining infrared sensors and underwater acoustics, under the aforementioned "small-world" assumptions (Ferrari and Wettergren, 2021). Optimal control and computational geometry solution approaches, such as cell decomposition (Cai and Ferrari, 2009), disjunctive programming (Swingler and Ferrari, 2013), and information roadmap methods (IRM) (Zhang et al., 2009), have been developed for optimizing robot performance by minimizing the cost of traveling through the workspace and processing sensor measurements, while maximizing the sensor rewards such as information gain. All these existing methods assume prior knowledge of sensor performance and of the workspace, and are applicable when the time and energy allotted to the robot are adequate for completing the sensing task. Information-driven path planning algorithm integrated with online mapping, developed in Zhu et al. (2019); Liu et al. (2019); Ge et al. (2011), have extended former treasure hunt solutions to problems in which a prior model of the workspace is not available and must be obtained online. Optimization-based algorithms have also been developed for fixed end-time problems with partial knowledge of the workspace, on the basis of the assumption that a probabilistic model of the information states and unlimited sensor measurements are available (Rossello et al., 2021). This paper builds on this previous work to develop heuristic strategies applicable when uncertainties cannot be learned or mathematically modeled in closed form, and the presence of external pressures might prevent task completion, e.g., adverse weather or insufficient time/energy.

Inspired by previous findings on human satisficing heuristic strategies (Gigerenzer and Gaissmaier, 2011; Gigerenzer, 1991; Gigerenzer and Goldstein, 1996; Gigerenzer, 2007; Oh et al., 2016), this paper develops, implements, and compares the performance between existing treasure hunt algorithms and human participants engaged in the same sensing tasks and experimental conditions by using a new design approach. Subsequently, human strategies and heuristics outperforming existing state-of-the-art algorithms are identified and modeled from data in a manner that can be extended to any sensor-equipped autonomous robot. The effectiveness of these strategies is then demonstrated with camera-equipped robots via high-fidelity simulations as well as physical laboratory experiments. In particular, human heuristics are modeled by using the "three building blocks" structure for formalizing general inferential heuristic strategies presented in Gigerenzer and Todd (1999). The mathematical properties of heuristics characterized by this approach are then compared with logic and statistics, according to the rationale in Gigerenzer and Gaissmaier (2011).

Three main classes of human heuristics for inferential decisions exist: recognition-based decision-making (Ratcliff and McKoon, 1989; Goldstein and Gigerenzer, 2002), one-reason decision-making (Gigerenzer, 2007; Newell and Shanks, 2003), and trade-off heuristics (Lichtman, 2008). Although categorized by respective decision mechanisms, these classes of human heuristics have been investigated in disparate satisficing settings, thus complicating the determination of which strategies are best equipped to handle different environmental pressures. Furthermore, existing human studies are typically confined to desktop simulations and do not account for action decisions pertaining to physical motion and path planning in complex workspaces. Therefore, this paper presents a new experimental design approach (Section 3) and tests in human participants to analyze and model satisficing active perception strategies (Section 7) that are generalizable and applicable to robot applications, as shown in Section 8.

The paper also presents new analysis and modeling studies of human satisficing strategies in both passive and active perception and decision-making tasks (Section 3). For passive tasks, time pressure on inference is introduced to examine subsequent effects on human decision-making in terms of decision model complexity and information gain. The resulting heuristic strategies (Section 5) extracted from human data demonstrate adaptability to varying time pressure, thus enabling inferential decision-making to meet decision deadlines. These heuristics significantly reduce the complexity of target feature search from an exhaustive search $O(2^n)$ to $O(nlog(n)+n)$, where n is the number of target features. Additionally, they exhibit superior classification performance when compared to optimizing strategies that utilize all target features for inference (Section 6), demonstrating the less-can-be-more effect (Gigerenzer and Gaissmaier, 2011).

For active tasks, when the sensing capabilities are significantly hindered, such as in adverse weather conditions, human strategies are found to amount to highly effective heuristics that can be modeled as shown in Section 7, and generalized to robots as shown in Section 8. The human strategies discovered from human studies are implemented on autonomous robots equipped with vision sensors and compared with existing planning methods (Section 8) through simulations and physical experiments in which optimizing strategies fail to complete the task or exhibit very poor performance. Under information cost pressure, a decision-making strategy developed using mixed integer nonlinear program (MINLP) (Cai and Ferrari, 2009; Zhang et al., 2009) was found to outperform existing solutions as well as human strategies (Section 8). By complementing the aforementioned heuristics, the MINLP optimizing strategies provide a toolbox for active robot perception under pressures that is verified both in experiments and simulations.

2 Treasure hunt problem formulation

This paper considers the active perception problem known as treasure hunt, in which a mobile information-gathering agent, such as a human or an autonomous robot, must find and localize all important targets, referred to as *treasures*, in an unknown workspace $\mathcal{W} \subset \mathbb{R}^3$. The number of possible treasures or targets, r, is unknown *a priori*, and each target i may constitute a treasure or another object, such as a clutter or false alarm, such that its classification may be represented by a random and discrete hypothesis variable Y_i with finite range $\mathcal{Y} = \{y_j \,|\, j \in \mathcal{J}\}$, where y_j represents the jth category of Y_i. While Y_i is hidden or non-observable, it may be inferred from $p_i \in \mathbb{Z}$ observed features among a set of n discrete random variables $X_i = \{X_{i,1}, \ldots, X_{i,n}\}$, and the lth $(1 \le l \le n)$ feature has a finite range $\mathcal{X}_l = \{x_{l,j} \,|\, j \in \mathcal{N}\}$ [see (Ferrari and Cai, 2009; Zhang et al., 2011; Ferrari and Vaghi, 2006) for more details]. At the onset of the search, X_i and Y_i are assumed unknown for all targets, as are the number of targets and treasures present in $\mathcal{W}$. Thus, the agent must first navigate the

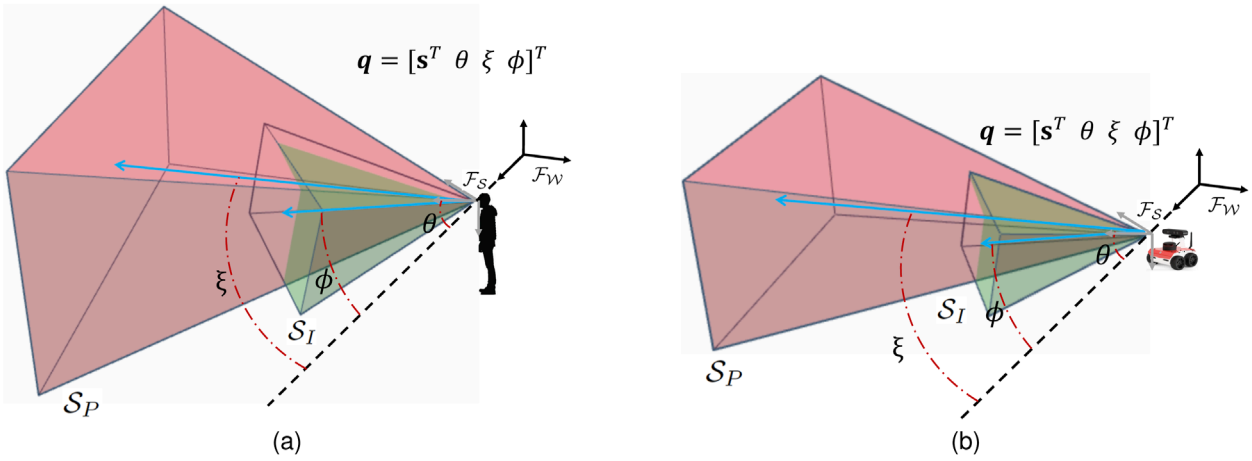


FIGURE 1
Human **(A)** and robot **(B)** state, configuration, and passive and active sensor FOVs.

workspace to find the targets and, then, observe their features to infer their classification.

All r targets are fixed, at unknown positions $\mathbf{x}_1, \ldots, \mathbf{x}_r \in \mathcal{W}$, and must be detected, observed, and classified using onboard sensors with bounded field-of-view (FOV) (Ferrari and Wettergren, 2021):

Definition 2.1: (Field-of-view (FOV)) For a sensor characterized by a dynamic state, in a workspace $\mathcal{W} \subset \mathbb{R}^3$, the FOV is defined as a closed and bounded subset $\mathcal{S} \subset \mathcal{W}$ such that a target feature $X_{i,l}$ may be observed at any point $\mathbf{x}_i \in \mathcal{S}$.

In order to obtain generalizable strategies for camera-equipped robots, in both human and robot studies knowledge of the targets is acquired, at a cost, through vision, and the sensing process is modeled by a probabilistic Bayesian network learned from data (Ferrari and Wettergren, 2021).

Although the approach can be easily extended to other sensor configurations, in this paper it is assumed that the information-gathering agent is equipped with one passive sensor for obstacle/target collision avoidance and localization, with FOV denoted by $\mathcal{S}_P$, and one active sensor for target inference and classification, with FOV denoted by $\mathcal{S}_I$ (Figure 1B). In human studies, the same passive/active configuration is implemented via virtual reality (VR) wand/joystick and goggles, and by measuring and constraining the human FOV, as shown in Figure 1A. Furthermore, the workspace is populated with q known fixed, rigid, and opaque objects $\mathcal{B}_1, \ldots, \mathcal{B}_q \subset \mathcal{W}$ that constitute obstacles as well as occlusions. Therefore, in order to observe the targets, the agent must navigate in $\mathcal{W}$ avoiding both collisions and occluded views, according to the following line of sight (LOS) visibility constraint:

Definition 2.2: (Line of sight) Given the sensor position $\mathbf{s} \in \mathcal{W}$, a target at $\mathbf{x} \in \mathcal{W}$ is occluded by an object $\mathcal{B} \subset \mathcal{W}$ if and only if,

$$L(\mathbf{s}, \mathbf{x}) \cap \mathcal{B} \neq \emptyset$$

where $L(\mathbf{s}, \mathbf{x}) = \{(1-\gamma)\mathbf{s} + \gamma\mathbf{x} \,|\, \gamma \in [0,1]\}$.

Let $\mathcal{F}_\mathcal{W}$ denote an inertial frame embedded in $\mathcal{W}$, and $\mathcal{I}$ denote the geometry of the agent body. The motion of the agent relative to the workspace can then be described by the position and orientation of a body frame $\mathcal{F}_\mathcal{S}$, embedded in the agent, relative to $\mathcal{F}_\mathcal{W}$. Thus, the state of the information-gathering agent at t_k can be described by the vector $\mathbf{q}_k = [\mathbf{s}_k^T \ \ \theta_k \ \ \xi_k \ \ \phi_k]^T$, where $\mathbf{s}_k$ represents the inertial position of the information-gathering agent in $\mathcal{W}$, $\theta_k \in \mathbb{S}^1$ is the orientation of the agent, and $\xi_k \in [\xi_l, \xi_u]$ and $\phi_k \in [\phi_l, \phi_u]$ are preferred sensing directions of the "passive" and "active" FOVs, respectively. In addition, ξ_l, ξ_u and ϕ_l, ϕ_u bound the preferred sensing directions for $\mathcal{S}_P$ and $\mathcal{S}_I$ with respect to the information-gathering agent body. By this approach it is possible to model FOVs able to move with respect to the agent body, as required by the motion of the human head or pan-tilt-zoom cameras (Figure 1).

Obstacle avoidance is accomplished by ensuring that the agent configuration, defined as $\mathbf{t}_k = [\mathbf{s}_k^T \ \ \theta_k]^T$, remains in free configuration space at all times. Let $\mathcal{C}$ represent all possible agent configurations, and $\mathcal{CB}_j = \{\mathbf{t} \in \mathcal{C} | \mathcal{I}(\mathbf{t}) \cap \mathcal{B}_j \neq \emptyset\}$ denote the C-obstacle associated with object $\mathcal{B}_j$ [defined in Ferrari and Wettergren (2021) and references therein). Then, the free configuration space is the space of configurations that avoid collisions with the obstacles or, in other words, that are the complement of all C-obstacle regions in $\mathcal{C}$, i.e., $\mathcal{C}_{\text{free}} = \{\mathcal{C} \backslash \bigcup_{j=1}^{q} \mathcal{CB}_j\}$.

According to directional visibility theory (Gemerek et al., 2022), the subset of the free space at which a target is visible by a sensor in the presence of occlusions can defined as follows:

Definition 2.3: (Target Visibility Region) For a sensor with FOV $\mathcal{S}_P \subset \mathcal{W}$, in the presence of q occlusions $\mathcal{B}_j (j = 1, \ldots, q)$ a target at $\mathbf{x}_i \in \mathcal{W}$ is visible within the target visibility region that satisfies both FOV and LOS conditions, i.e.,:

$$\mathcal{TV}_i = \left\{\mathbf{t} \in \mathcal{C}_{\text{free}} \mid \mathbf{x}_i \in \mathcal{S}_P, L(\mathbf{s}, \mathbf{x}_i) \cap \mathcal{B}_j = \emptyset, \forall j\right\}$$

It follows that multiple targets are visible to the sensor in the intersection of multiple visibility regions defined as Gemerek et al. (2022):

Definition 2.4: (Set Visibility Region) Given a set of r target-visibility regions $\{\mathcal{TV}_i \mid i \in \{1, 2, \ldots, r\}\}$, let $S \subseteq \{1, 2, \ldots, r\}$ represent

the set of target indices of two or more intersecting regions, such that the following holds $\bigcap_{i\in S}\mathcal{TV}_i \neq \emptyset$. Then, the set visibility region of target i is defined as

$$\mathcal{V}_S = \left\{ \bigcap_{i\in S} \mathcal{TV}_i \mid S \subseteq \{1,2,\ldots,r\} \right\}$$

Similarly, after a target i is detected and localized, the agent may observe the target features using the active sensor with FOV $\mathcal{S}_I$ provided $\mathbf{x}_i \in \mathcal{S}_I(\mathbf{q})$ and $L(\mathbf{s},\mathbf{x}_i) \cap \mathcal{B}_j = \emptyset$, $1 \le j \le q$. In order to explore the tradeoff of information value and information cost in inferential decisions, use of the active sensor is associated with an information cost $J(t_k)$ that may reflect the use of processing power, data storage, and/or need for covertness. Then, the information-gathering agent, must make a deliberate decision to observe one or more target features prior to obtaining the corresponding measurement, which may consist of an image or raw measurement data from which feature X_i may be extracted. For simplicity, measurement errors are assumed negligible but they may be easily introduced following the approach in [10, Chapter 9]. Then, the goal of the treasure hunt is to infer the hypothesis variable Y_i from X_i, $i = 1,2,\ldots$, using a probabilistic measurement model $P(Y_i,X_{i,1},\ldots,X_{i,n})$ (Ferrari and Cai, 2009). The measurement model, chosen here as a Bayesian network (BN) (Figure 2D), consists of a probabilistic representation of the relationship between the observed target features and the target classification that may be learned from expert knowledge or prior training data as shown in [10, Chapter 9]. Importantly, because the agent may not have the time and/or resources to observe all target features, classification may be performed from a sequence of partial observations.

Target features are observed through test decisions made by the information-gathering agent, which result into soft or hard evidence for the probabilistic model $P(Y_i,X_{i,1},\ldots,X_{i,n})$ (Jensen and Nielsen, 2007). Let $u(t_k) \in \mathcal{U}_k$ denote at time t_k test decision chosen from the set of all admissible tests $\mathcal{U}_k \subset \mathcal{U}$. The set $\mathcal{U} = \{\vartheta_c, \vartheta_s, \vartheta_{un}\}$ consists of all test decisions, where ϑ_c and ϑ_s represent the decisions to continue or stop observing target features, and ϑ_{un} represents the decision to not observe any feature. The test decision $u(t_k)$ generates a measurement variable at time step t_{k+1},

$$z(t_{k+1}) = x_{i,l}, \quad 1 \le i \le r, \quad 1 \le l \le n, \quad x_{i,l} \in \mathcal{X}_l$$

observed after paying the information cost $J(t_k) \in \mathbb{Z}$, which is modeled as cumulative number of observed features up to t_k. When the measurement budget R is finite, it may not be exceeded by the agent and, thus, the treasure hunt problem must be solved subject to the hard constraint

$$J(t_k) \le R.$$

Action decisions modify the state of the world and/or information-gathering agent (Jensen and Nielsen, 2007). In the treasure hunt problem, action decisions are control inputs that decide the position and orientation of the agent and of the FOVs $\mathcal{S}_P$ and $\mathcal{S}_I$. Let $a(t_k) \in \mathcal{A}_k$ denote an action decision chosen at time t_k from set $\mathcal{A}_k$ of all admissible actions. The agent motion can then be described by a causal model as the following difference equation,

$$\mathbf{q}_{k+1} = \mathbf{f}[\mathbf{q}_k, a(t_k), t_k]$$

where $\mathbf{f}[\cdot]$ is obtained by modeling the agent dynamics.

Then, an active perception strategy consists of a sequence of action and test decisions that allow the agent to search the workspace and obtain measurements from targets distributed therein, as follows:

Definition 2.5: (Inferential Decision Strategy) An active inferential decision strategy is a class of admissible policies that consists of a sequence of functions,

$$\sigma = \{\pi_0, \pi_1, \ldots, \pi_T\}$$

where $\pi_k$maps all past information-gathering agent states, test variables, action and test decisions into admissible action and test decisions,

$$\{a(t_k), u(t_k)\} = \pi_k[\mathbf{q}_0, a(t_1), u(t_1), z(t_1), J(t_1), \mathbf{q}_1, \ldots, a(t_{k-1}), u(t_{k-1}), z(t_{k-1}), J(t_{k-1}), \mathbf{q}_{k-1}]$$

such that $\pi_k[\cdot] \in \{\mathcal{A}_k, \mathcal{U}_k\}$, for all $k = 1,2,\ldots,T$.

Based on all the aforementioned definitions, the problem is formulated as follows:

Problem 1: (Satisficing Treasure Hunt)

Given an initial state $\mathbf{q}_0$ and the satisificing aspiration level of total information value Δ, the satisficing treasure hunt problem consists of finding an active inferential decision making strategy, σ, over a known and finite time horizon $(0,T]$, such that the cumulative information value collected from all observed features is no less than Δ,

$$\sum_{i=1}^{r} \left[1\left(\exists k, \mathbf{x}_i \in \mathcal{S}_I(\mathbf{q}_k) \wedge L(\mathbf{s}_k, \mathbf{x}_i) \cap \mathcal{B}_j, \forall j\right) I(Y_i; X_i) \right] \ge \Delta \tag{1}$$

where

$$\mathbf{q}_{k+1} = \mathbf{f}[\mathbf{q}_k, a(t_k), t_k] \tag{2}$$

$$\hat{y}_i = \arg\max_{y\in\mathcal{Y}} P(Y_i = y, X_{i,1}, \ldots, X_{i,n}) \tag{3}$$

$$I(Y_i; X_i) = H(Y_i) - H(Y_i \mid X_i) \tag{4}$$

$$J(t_T) \le R \tag{5}$$

$$i = 1,2,\ldots,r, \; 1 \le k \le T \tag{6}$$

$$j = 1,2,\ldots,q \tag{7}$$

An optimal search strategy makes use of the agent motion model (Equation 2), measurement model (Equation 3) and knowledge of the workspace $\mathcal{W}$ to maximize the information value while minimizing the distance traveled and the cumulative information cost (Ferrari and Wettergren, 2021). A feasible search strategy may use all or part of the available models of the environment and targets, or knowledge of prior states and decisions to produce a sequence of action and test decisions that satisfy the objective (Equation 1) by the desired end time t_T.

3 Human satisficing studies

Human strategies and heuristics for active perception are modeled and investigated by considering two classes of satisficing

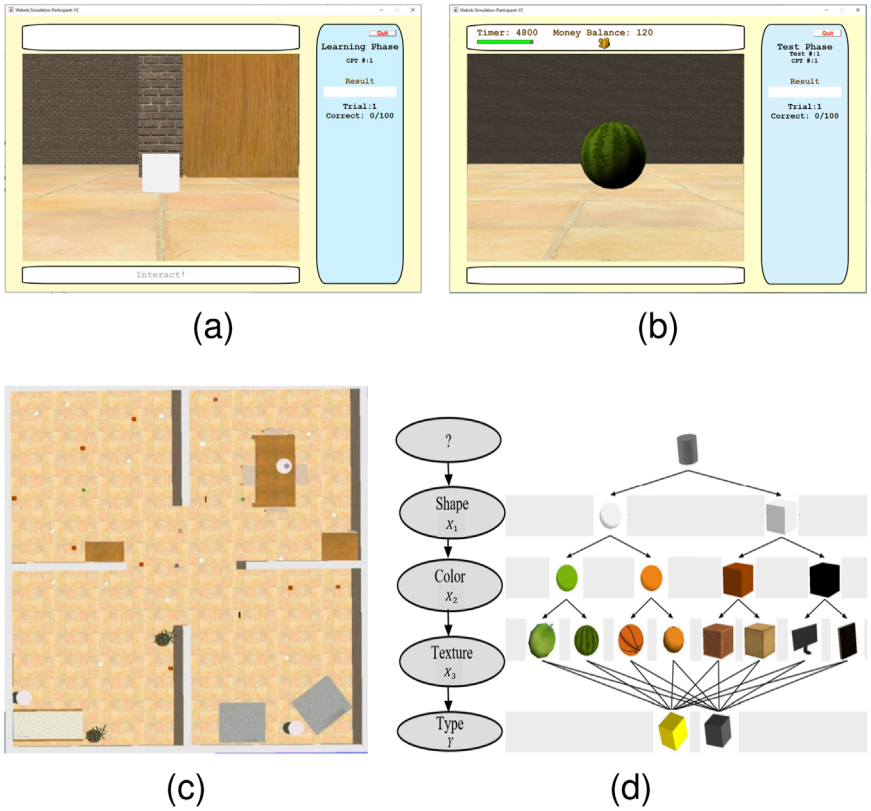


FIGURE 2
First-person view in training phase without prior target feature revealed **(A)** and with feature revealed by a participant **(B)** in the Webots® workspace **(C)** and target features encoded in a BN structure with ordering constraints **(D)**.

treasure hunt problems, referred to as passive and active experiments. Passive satisficing experiments focus on treasure hunt problems in which information is presented to the decision maker who passively observes features needed to make inferential decisions. Active satisficing experiments allow the decision maker to control the amount of information gathered in support of inferential decisions. Additionally, treasure hunt problems with both static and dynamic robots are considered in order to compare with and extend previous satisficing studies, evolving human studies traditionally conducted on a desktop (Oh et al., 2016; Toader et al., 2019; Oh-Descher et al., 2017) to ambulatory human studies in virtual reality that parallel mobile robots applications (Zielinski et al., 2013).

Previous cognitive psychology studies showed that the urgency to respond (Cisek et al., 2009) and the need for fast decision-making (Oh et al., 2016) significantly affect human decision evidence accumulation, thus leading to the use of heuristics in solving complex problems. Passive satisficing experiments focus on test decisions, which determine the evidence accumulation of the agent based on partial information under "urgency". Inspired by satisficing searches for Spanish treasures with feature ordering constraints (Simon and Kadane, 1975), active satisficing includes both test and action decisions, which change not only the agent's knowledge and information about the world but also its physical state. Because information gathering by a physical agent such as a human or robot is a causal process (Ferrari and Wettergren, 2021), feature ordering constraints are necessary in order to describe the temporal nature of information discovery.

Both passive and active satisficing human experiments comprise a training phase and a test phase that are also similarly applied in the robot experiments in Sections 6–8. During the training phase, human participants learn the validity of target features in determining the outcome of the hypothesis variable. They receive feedback on their inferential decisions to aid in their learning process. During the test phase, pressures are introduced, and action decisions are added for active tasks. Importantly, during the test phase, no performance feedback or ground truth is provided to human participants (or robots).

3.1 Passive satisficing task

The passive satisficing experiments presented in this paper adopted the passive treasure hunt problem, shown in Supplementary Figure S2 and related to the well-known weather prediction task (Gluck et al., 2002; Lagnado et al., 2006; Speekenbrink et al., 2010). The problem was first proposed in Oh et al. (2016) to investigate the cognitive processes involved in human test decisions under pressure. In view of its passive nature, the experimental platform of choice consisted of a desktop computer used to emulate the high-paced decision scenarios, and to encourage

the human participants to focus on cue(feature) combination rather than memorization (Oh et al., 2016; Lamberts, 1995).

The stimuli presented on a screen were precisely controlled, ensuring consistency across participants and minimizing distractions from irrelevant objects or external factors (Garlan et al., 2002; Lavie, 2010). In each task, participants were presented with two different stimuli from which to select the "treasure" before the total time, t_T, at one's disposal has elapsed (time pressure). The treasures are hidden but correlated with the visual appearance of the stimulus, and the underlying probabilities must be learned by trial and error during the training phase. Each stimulus is characterized by four binary cues or "features", namely, color (X_1), shape (X_2), contour (X_3), and line orientation (X_4), illustrated in the table in Supplementary Figure S2. The goal of this passive satisficing task is to find all treasures among stimuli that are presented on the screen or, in other words, to infer a binary hypothesis variable Y, with range $\mathcal{Y} = \{y_1, y_2\}$, where y_1 = "treasure" and y_2 = "not treasure". The task is passive by design because the participant cannot control the information displayed in order to aid his/her decisions.

During the training phase, each (human) participant performed 240 trials in order to learn the relationship between features, $X = \{X_1, X_2, X_3, X_4\}$, and the hypothesis variable Y. After the training phase, participants were divided into two groups. The first group underwent a moderate time pressure (TP) experiment and was tested against two datasets, each consisting of 120 trials. Participants were required to make decisions within a response time t_T = 750 ms, which allowed ample time to ponder on the features presented and how they related to the treasure. The second group underwent an intense TP experiment, with a response time of only t_T = 500 ms. Participants in this group also encountered two datasets, each containing 120 trials. A more detailed description of the experiment, including redundant features, and human subject procedures that informed, among other parameters, the number of trials can be found in Oh et al. (2016). Subsequently, the task was modified to develop a number of active satisficing treasure hunts in which information about the treasures had to be obtained by navigating a complex environment, as explained in the next section.

As shown in Table 1, the relevant statistics for passivesatisficing experiments are summarized in the upper part. Similarly, the statistics for active satisficing experiments are presented in the lower part, where the human participants are allowed to move in an environment and choose the interaction order with the targets. The statistics correspond to three conditions: "No Pressure", "Info Cost Pressure", and "Sensory Deprivation". These pressure conditions will be introduced in detail in Section 3.2.

3.2 Active satisficing treasure hunt task

The satisficing treasure hunt task is an ambulatory study in which participants must navigate a complex environment populated with a number of obstacles and objects in order to first find a set of targets (stimuli) and, then, determine which are the treasures. Additionally, once the targets are inside the participant's FOV, features are displayed sequentially to him/her only after paying cost for the information requested. The ordering constraints (illustrated in Figure 2D) allow for the study of information cost and its role in the decision making process by which the task is to be performed not only under time pressure but also a fixed budget. Thus, the satisficing treasure hunt allows not only to investigate how information about a hidden variable (treasure) is leveraged, but also how humans mediate between multiple objectives such as obstacle avoidance, limited sensing resources, and time constraints. Participants must, therefore, search and locate the treasures without any prior information on initial target features, target positions, or workspace and obstacle layout.

In order to utilize a controlled environment that can be easily changed to study all combinations of features, target/obstacle distributions, and underlying probabilities, the active satisficing treasure hunt task was developed and conducted in a virtual reality environment known as the DiVE (Zielinski et al., 2013). By this approach different experiments were designed and easily modified so as to investigate different difficulty levels and provide the human participants repeatable, well-controlled, and immersive experience of acquiring and processing information to generate behavior (Van Veen et al., 1998; Pan and Hamilton, 2018; Servotte et al., 2020). The DiVE consists of a 3 m × 3 m × 3 m stereoscopic rear projected room with head and hand tracking, allowing participants to interact with a virtual environment in real-time (Zielinski et al., 2013). By developing a new interface between the DiVE and the robotic software Webots®, this research was able to readily introduce humans within the same environments designed for humans, and *vice versa*, according to the BN model of the desired treasure hunt task. The structure of the BN used for the human/robot treasure hunt perception task is plotted in Figure 2D. The BN parameters, not shown for brevity, were varied across trials to obtain a representative dataset from the human study from which mathematical models of human decision strategies could be learned and validated.

Six human participants were trained and given access to the DiVE for a total of fifty-four trials with the objective to model aspects of human intelligence that outperform existing robot strategies. The number of trials and participants is adequate to the scope of the study which was not to learn from a representative sample of the human population, but to extract inferential decision making strategies generalizable to treasure hunt robot problems. Besides manageable in view of the high costs and logistical challenges associated with running DiVE experiments, the size of the resulting dataset was also found to be adequate to varying all of the workspace and target characteristics across experiments, similarly to the studies in Ziebart et al. (2008); Levine et al. (2011). Moreover, through the VR googles and environment, it was possible to have precise and controllable ground truth not only about the workspace, but also about the human FOV, $\mathcal{S}_p$, within which the human could observe critical information such as targets, features, and obstacles.

A mental model of the relationship between target features and classification was first learned by the human participants during 100 stationary training sessions (Figures 2A, B) in which the target features (visual cues), comprised of shape (X_1), color (X_2), and texture (X_3), followed by the target classification Y, where $\mathcal{Y} = \{y_1, y_2\}$, were displayed on a computer screen, through the desktop Webots® simulation shown in Figure 2. Participants were then instructed to search for treasures inside an unknown 10 m × 10 m Webots® workspace with $r = 30$ targets (Figure 2C), by paying information cost $J(t_k)$ to see the features, $X_i = \{X_{i,1}, X_{i,2}, X_{i,3}\}$, of every

TABLE 1 Experiment conditions and trials.

Experiment type	Pressure condition	Number of participants	Number of training targets for each participant	Number of test trials for each participant
Passive Satisficing	NP	48	240	120
	Moderate TP	48	240	120
	Intense TP	48	240	120
Active Satisficing	No Pressure	6	100	3
	Info Cost Pressure	6	100	3
	Sensory Deprivation	6	100	3

target (labeled by i) inside their FOV sequentially over time (test phase). Based on the features observed, which may have included one or more features in the set X, participants were asked to decide which targets were treasures ($Y = y_1$) or not ($Y = y_2$). No feedback about their decisions was provided and, as explained in Section 2, the task had to be performed within a limited budget R and time period t_T.

Mobility and ordering feature constraints are both critical to autonomous sensors and robots, because they are intrinsic to how these cyber-physical systems gather information and interact with the world around them. Thanks to the simulation environments and human experiment design presented in this section, we were able to engage participants in a series of classification tasks in which target features were revealed only after paying both a monetary and time cost, similarly to artificial sensors that require both computing and time resources to process visual data. Participants were able to build a mental model built for decision making with the inclusion of temporal constraints during the training phase, according to the BN conditional probabilities (parameters) of each study. By sampling the Webots® environments from each BN model, selected by the experiment designer to encompass the full range of inference problem difficulty, and by transferring them automatically into VR (Figure 3) the data collected was guaranteed ideally suited for the modeling and generalization of human strategies to robots (Section 7). As explained in the next section, the test phase was conducted under three conditions: no pressure, money pressure, and sensory deprivation (fog).

4 External pressures inducing satisficing

Previous work on human satisficing strategies and heuristics illustrated that most humans resort to these approaches for two main reasons, one is computational feasibility and the other is the "less-can-be-more" effect (Gigerenzer and Gaissmaier, 2011). When the search for information and computation costs become impractical for making a truly "rational" decision, satisficing strategies adaptively drop information sources or partially explore decision tree branches, thus accommodating the limitations of computational capacity. In situations in which models have significant deviations from the ground truth, external uncertainties are substantial, or closed-form mathematical descriptions are lacking, optimization on potentially inaccurate models can be risky. As a result, satisficing strategies and heuristics often outperform classical models by utilizing less information. This effect can be explained in two ways. Firstly, the success of heuristics is often dependent on the environment. For example, empirical evidence suggests that strategies such as "take-the-best," which rely on a single good reason, perform better than classical approaches under high uncertainty (Hogarth and Karelaia, 2007). Secondly, decision-making systems should consider trade-offs between bias and variance, which is determined by model complexity (Bishop and Nasrabadi, 2006). Simple heuristics with fewer free parameters have smaller variance than complex statistical models, thus avoiding overfitting to noisy or unrepresentative data, and generalizable across a wider range of datasets (Bishop and Nasrabadi, 2006; Brighton et al., 2008; Gigerenzer and Brighton, 2009).

Motivated by the situations where robots' mission goals can be severely hindered or completely compromised due to inaccurate environment or sensing models caused by pressures, the paper seeks to emulate aspects of human intelligence under the pressures and study their influence on decisions. The environment pressures include, for example, time pressure (Payne et al., 1988), information cost (Dieckmann and Rieskamp, 2007; Bröder, 2003), cue(feature) redundancy (Dieckmann and Rieskamp, 2007; Rieskamp and Otto, 2006), sensory deprivation, and high risks (Slovic et al., 2005; Porcelli and Delgado, 2017). Cue(feature) redundancy and high risk have been investigated extensively in statistics and economics, particularly in the context of inferential decisions (Kruschke, 2010; Mullainathan and Thaler, 2000). In the treasure hunt problem, sensory deprivation and information cost directly and indirectly influence action decisions, which brings insight how these pressures impact agents' motion. However, the effects of sensory deprivation on human decisions have not been thoroughly investigated compared to other pressures. Time pressure is ubiquitous in the real world, yet heuristic strategies derived from human behavior are still lacking. Thus, this paper aims to fill this research gap by examining the time pressure, information

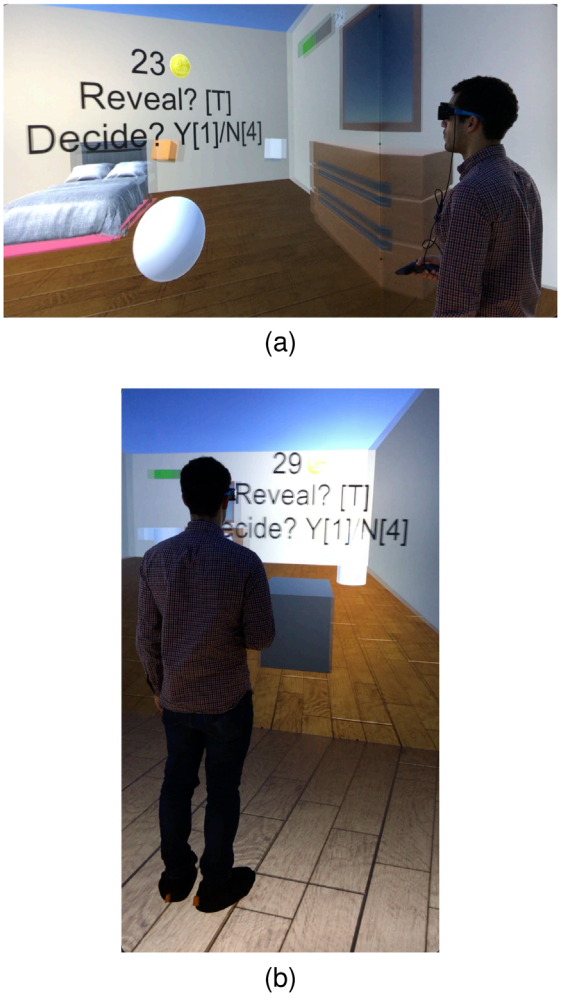


FIGURE 3
Test phase in active satisficing experiment in DiVE from side view **(A)**, and from rear view **(B)**.

cost pressure, and sensory deprivation and their effects on decision outcomes.

4.1 Time pressure

Assume that a fixed time interval t_c is needed to integrate one additional feature into the inference decision-making process. In the meantime, each decision must be made within t_T, and p_i is the number of observed features for the ith target. The satisficing strategies must adaptively select a subset of the features such that a decision is made within the time constraint

$$p_i t_c < t_T, \quad i = 1, 2, \ldots, r$$

According to the human studies in Oh et al. (2016), the response time of participants in the passive satisficing tasks was measured during the pilot work. The average response time in these tasks was found to be approximately 700 ms. Based on this finding, three time windows were designed to represent different time pressure levels: a 2-s time window was considered without any time pressure; a 750 ms time window was considered moderate time pressure; and a 500 ms time window was considered intense time pressure.

4.2 Information cost

The cost of acquiring new information intrinsically makes an agent use fewer features to reach a decision. In Section 2, new information for the ith target is collected through a sequence of p_i observed target features. Thus, for all r targets, the information cost is mathematically described as the total number of observed features not exceeding a preset budget R

$$\sum_{i=1}^{r} p_i \leq R$$

In Section 3.2, the human studies introduce information cost pressure using the parameter $R = 30$. In the context of the treasure hunt problem, R represents the measurement budget, which limits the number of features that a participant can observe from targets. In this experiment, for example, a total of $r = 30$ targets was used, and an information budget of $R = 30$ was chosen such that the human participants were able to observe, on average, one feature per target. Other experiments were similarly performed by considering a range of parameters that spanned task difficulty levels across participants and treasure hunt types.

4.3 Sensory deprivation

As explained in Section 2, information-gathering agents were not provided a map of the workspace $\mathcal{W}$ *a priori*, and, instead, were required to obtain information about target and obstacle positions and geometries by means of a passive on-board sensor (e.g., camera or LIDAR) with FOV $\mathcal{S}_P$ as shown in Figure 4A. From the definition of set visibility region (Definition 2.4), for a subset $S \subseteq \{1, 2, \ldots, r\}$ of target indices, the set visibility region $\mathcal{V}_S \subseteq \mathcal{C}_{\text{free}}$ contains all targets in S visible to passive sensor with FOV $\mathcal{S}_P$. A globally optimal solution to treasure hunt problem (Equations 1–7) with respect to a subset of targets S is feasible if and only if $\mathcal{V}_S \neq \emptyset$.

In parallel to the human studies in Section 3.2, robot sensory deprivation was introduced by simulating/producing fog in the workspace, thereby reducing the FOV radius to approximately 1 m, in a 20 m × 20 m robot workspace. A fog environment is simulated inside the Webots® environment as shown in Figure 4, thereby reducing the camera's ability (Figure 4B) to view targets inside the sensor $\mathcal{S}_P$. As a result, $V_S = \emptyset$ even when there are $|S| = 2$ targets, indicating that a globally optimal solution is infeasible. Consequently, optimal strategies typically fail under sensory deprivation due to lack of target information. Using the methods presented in the next section, human strategies for modulating between satisficing and optimizing strategies are first learned from data and, then, generalized to autonomous robots, as shown in Section 8. Satisficing strategies are aimed at overcoming this difficulty, and use local information to explore the environment and visit targets.

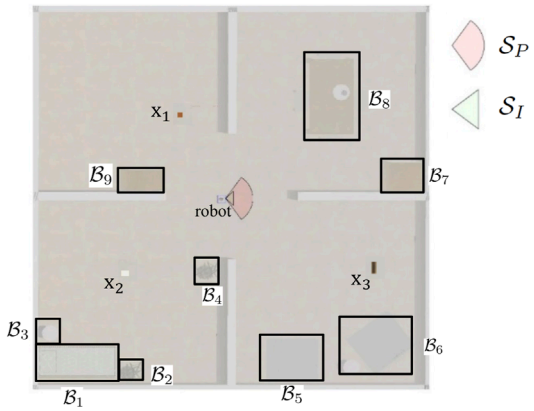


FIGURE 4
Top view visibility conditions of the unknown workspace **(A)** and first-person view of poor visibility condition **(B)** due to fog.

5 Mathematical modeling of human passive satisficing strategies

Previous work by the authors showed that human participants drop less informative features to meet pressing time deadlines that do not allow them to complete the tasks optimally (Oh et al., 2016). The analysis of data obtained from the moderate TP experiment (Figure 5A) and intense TP experiment (Figure 5B) reveals similar interesting findings regarding human decision-making under different time pressure conditions. Under the no TP condition, the most probable decision model selected by human participants (indicated by the yellow contour for D15 in Figures 5A, B) utilizes all four features and aims at maximizing information value. However, under moderate TP, the most probable decision model selected by human participants (indicated by a red box in Figure 5A) uses only three features and has lower information value than the no TP condition. As time pressure becomes the most stringent in the intense TP, the most probable decision model selected by human participants (indicated by a dark blue box in Figure 5B) uses only two features and exhibits even lower information value than observed in the previous two time pressure conditions. Figure 5C shows all possible decision models (i.e. features combinations) that a participant can use to make an inferential decision. These results demonstrate the trade-offs made by human participants among time pressure, model complexity, and information value. As time pressure increases, individuals adaptively opt for simpler decision models with fewer features, and sacrificed information value to meet the decision deadline, thus reflecting the cognitive adaptation of human participants in response to time constraints.

5.1 Passive satisficing decision heuristic propositions

Inspired by human participants' satisficing behavior indicated by the data analysis above, this paper develops three heuristic decision models, which accommodate varying levels of time pressure and adaptively select a subset of information-significant features to solve the inferential decision making problems. For simplicity and based on experimental evidence, it was assumed that observed features were error free.

5.1.1 Discounted cumulative probability gain (ProbGain)

The heuristic is designed to incorporate two aspects of behaviors observed from human data. First, the heuristic encourages the use of features that provide high information value for decision-making. By summing up the information value of each feature, the heuristic prioritizes the features that contribute the most to evidence accumulation. Second, the heuristic also considers the cost of using multiple features in terms of processing time. By applying a higher discount to models with more features, the heuristic discourages excessive cost on time that might lead to violation of time constraints.

For an inferential decision-making problem with sorted p observed features $\{x_j\}_{j=1}^{p}$ according to the information value $v_{\text{ProbGain}}(x_j)$ in descending order, where $v_{\text{ProbGain}}(x_j)$ representing the increase in information value with respect to the maximum a-posterior rule

$$v_{\text{ProbGain}}\left(x_j\right) = \max_{y\in\mathcal{Y}} p\left(Y=y \mid x_j\right) - \max_{y\in\mathcal{Y}} p\left(Y=y\right)$$

Let $\{x_1, x_2, \ldots, x_i\}$ represent a subset of observed features that contains the first (i) most informative features with respect to $v_{\text{ProbGain}}(x_j)$, where t_T is the allowable time to make a classification decision, and the discount factor $\gamma \in (0,1)$ is defined to be a function of t_T in order to represent the penalty induced by time pressure. Then, the heuristic strategy can be modeled as follows,

$$H_{\text{ProbGain}}\left(t_T, \{x_j\}_{j=1}^{p}\right) = \arg\max_i \left\{\gamma(t_T)^i \sum_{j=1}^{i} v_{\text{ProbGain}}\left(x_j\right)\right\}$$

where,

$$\gamma(t_T) = \exp\left(-\frac{\lambda}{t_T}\right)$$

and, thus, λ may be used to represent the extent to which the discount γ is applied to the cue information value.

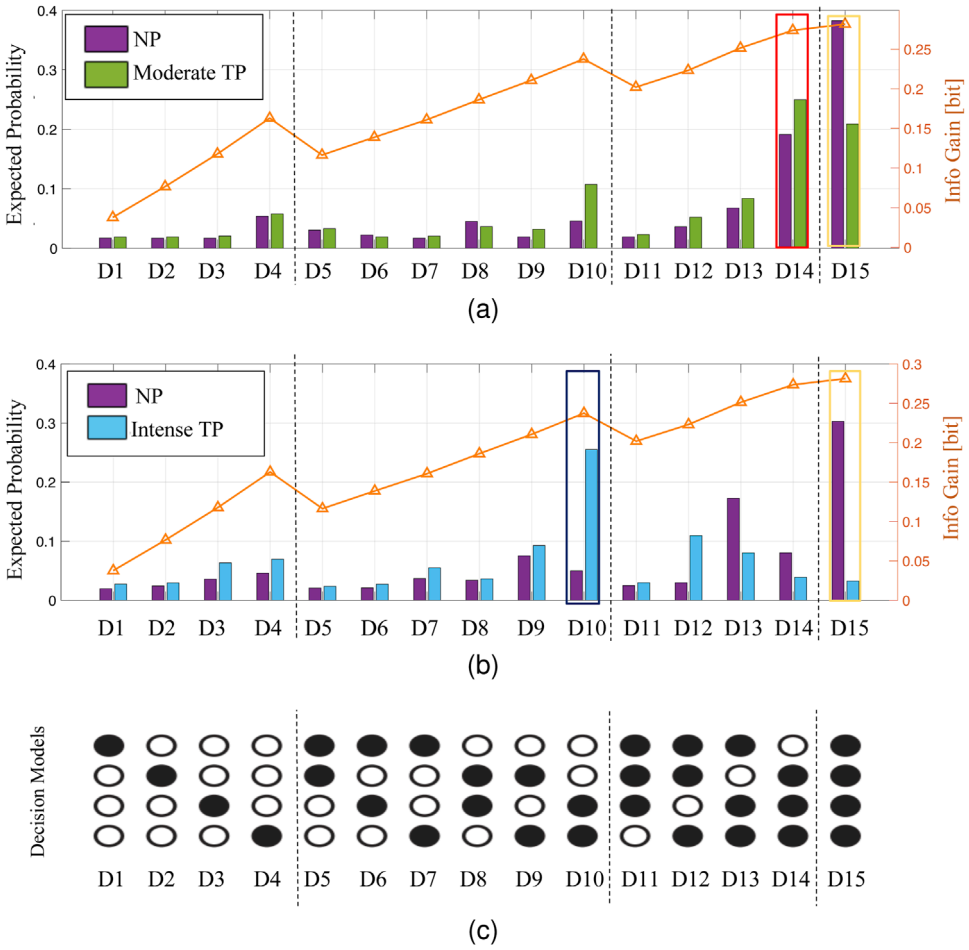


FIGURE 5
Human data analysis results for the moderate TP experiment **(A)** and the intense TP experiment **(B)** with the enumeration of decision models **(C)**.

5.1.2 Discounted log-odds ratio (LogOdds)

Log odds ratio plays a central role in classical algorithms like logistic regression (Bishop and Nasrabadi, 2006), and represents the "confidence" of making a inferential decision. The update of log odds ratio with respect to a "new feature" is through direct summation, thus taking advantage of the feature independence and arriving at fast evidence accumulation. Furthermore, the use of log odds ratio in the context of time pressure is slightly modified such that a discount is applied with inclusion of an additional feature to penalize the feature usage because of time pressure. By combining the benefits of direct summation for fast evidence accumulation and the discount for time pressure as inspired from human behavior, the heuristic based on log odds ratio can make efficient decisions by considering the most relevant features under time constraints.

For an inferential decision-making problem with sorted p observed features $\{x_j\}_{j=1}^{p}$ according to the information value $|v_{\text{ProbGain}}(x_j)|$ in descending order, where $|v_{\text{ProbGain}}(x_j)|$ represents the log odds ratio of observed features x_j. Then, the heuristic strategy can be modeled as follows,

$$H_{\text{LogOdds}}\left(t_T,\{x_j\}_{j=1}^{p}\right)=\arg\max_i\left\{\gamma(t_T)^i \mid v_0+\sum_{j=1}^{i} v_{\text{ProbGain}}\left(x_j\right) \mid\right\}$$

where

$$v_I(x_j)=log\left(p\left(x_j \mid y_1\right)\right)-log\left(p\left(x_j \mid y_2\right)\right)$$

$$v_0=log\left(p\left(Y=y_1\right)\right)-log\left(p\left(Y=y_2\right)\right)$$

5.1.3 Information free feature number discounting (InfoFree)

The previous two feature selection heuristics are both based on comparison: multiple candidate sets of features are evaluated and compared, and the heuristics select the one with the best trade-off between information value and processing time cost. A simpler heuristic is proposed to avoid comparisons and reduces the computation burden, while still showing the behavior that dropping less informative features due to time pressure observed from human participants.

Sort the p features according to the information value $v_I(x_j)$ in descending order as $x_1,x_2,\ldots,x_p$, and a subset of the first i most informative features refers to as $\{x_1,x_2,\ldots,x_i\}$. The heuristic strategy

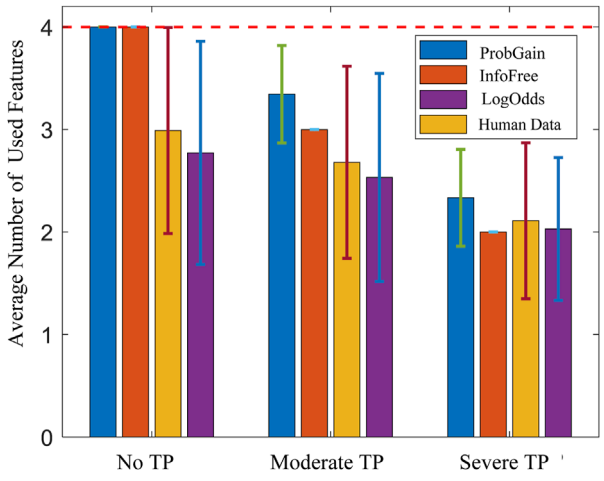


FIGURE 6
Mean and standard deviation of the number of used features of three heuristic strategies and the human strategy under three time pressure levels.

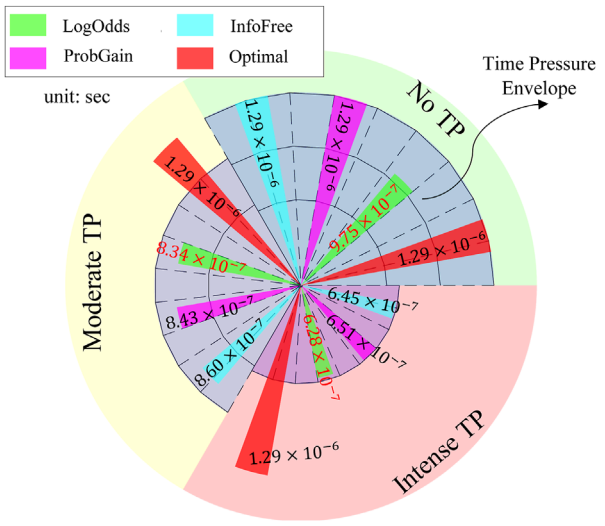


FIGURE 7
Processing time (unit: sec) of three time-adaptive heuristics and the "Bayes optimal" strategy.

is as follows

$$H_{\text{InfoFree}}(t_T) = \left\lceil p \exp\left(-\frac{\lambda}{t_T}\right)\right\rceil$$

The outputs of the three heuristics are the numbers of features to be fed into the model $P(Y_i, X_{i,1}, \ldots, X_{i,n})$ to make an inference decision. Some mathematical properties (e.g., convergence and monotonicity) of the three proposed heuristic strategies are presented in Supplementary Appendix SA1.

5.2 Model fit test against human data

The model fit tests against human data of the three proposed time-adaptive heuristics are under three time pressure levels, with the time constraints scaled to ensure comparability between human experiments and heuristic tests. The results, as shown in Figure 6, indicate two major observations. First, as time pressure increases, all three strategies utilize fewer features, thus demonstrating their adaptability to time constraints and mirroring the behavior observed in human participants. Second, among the three strategies, H_{LogOdds} exhibits the closest average number of features and standard deviation to the human data across all time pressure conditions. Consequently, H_{LogOdds} is the heuristic strategy that best matches the human data among the three proposed strategies.

6 Autonomous robot applications of passive satisficing strategies

The effectiveness of the human passive satisficing strategies modeled in the previous section, namely, the three heuristics denoted by H_{ProbGain}, H_{LogOdds}, and H_{InfoFree}, was tested on an autonomous robot making inferential decisions on the well-established database known as car evaluation dataset (Dua and Graff, 2017). This dataset, containing 1,728 samples, is chosen over other benchmark problems because its size is comparable to the database used for modeling human heuristics and is characterized by six possibly redundant features, which allows for the ability to adaptively select a subset of features to infer the target class. The performance of the three heuristics is compared against that of a naïve Bayes classifier, referred to as "Bayes optimal" herein, which utilizes all available features for decision-making.

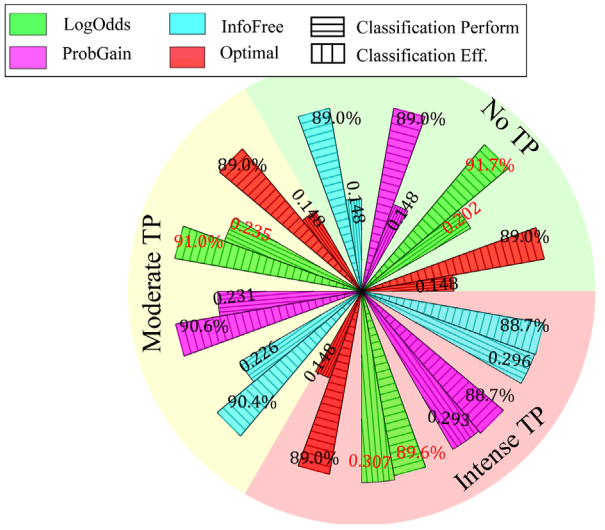


FIGURE 8
Classification performance and efficiency of three time-adaptive heuristics under three time pressure conditions.

The car evaluation dataset records the cars' acceptability, on the basis of six features and originally four classes. The four classes are merged into two. A training set of 1,228 samples is used to learn the conditional probability tables (CPTs), ensuring equal priors for

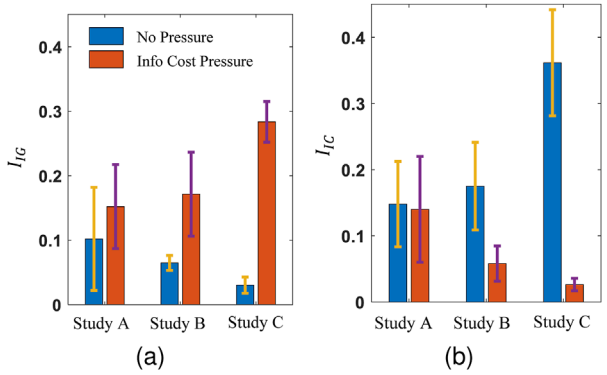


FIGURE 9
(A) The information value attempt index, I_{IG}, and **(B)** information-cost parsimony index I_{IC} are shown for high-performance human participants under two pressure conditions.

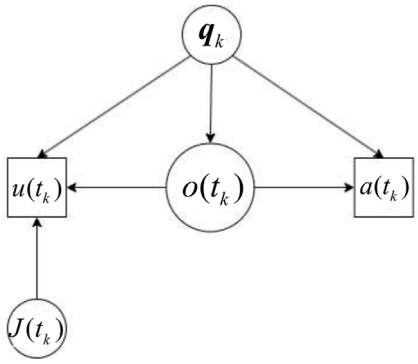


FIGURE 10
The intra-slice DBN that models human decision behavior.

both classes. After learning the CPTs, 500 samples are used to test the classification performance of the heuristics and the naïve Bayes classifier. The tests are conducted under three conditions: no TP, moderate TP, and intense TP.

The experiments are performed on a digital computer using MATLAB R2019b on an AMD Ryzen 9 3900X processor. The processing times of the strategies are depicted in Figure 7. If a heuristic's processing time falls within the time pressure envelope (blue area), the time constraints are considered satisfied. The no TP condition provides sufficient time for all heuristics to utilize all features for decision-making. The moderate TP condition allows for 75% of the time available in the no TP condition, whereas the intense TP condition allows for 50% of the time available in the no TP condition. All three heuristics are observed to satisfy the time constraints across all time pressure conditions.

The classification performance and efficiency of the three time-adaptive strategies is plotted in Figure 8. H_{LogOdds} outperforms the other three strategies on this dataset, and its performance

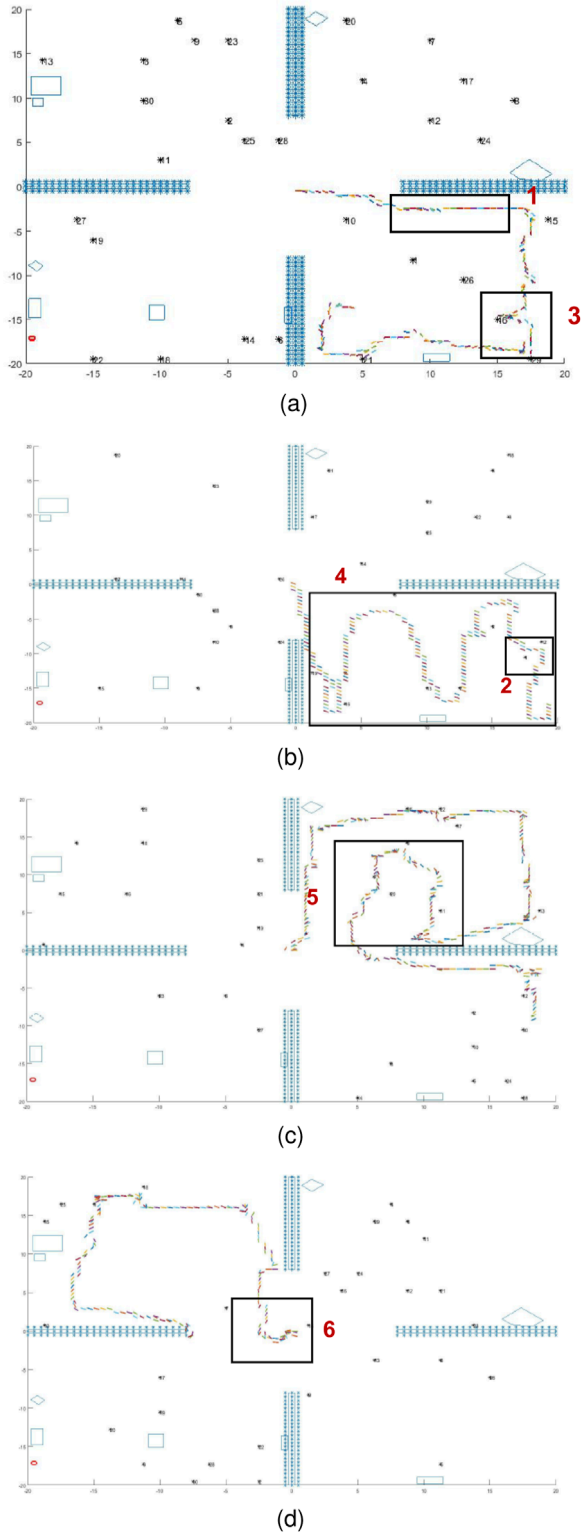


FIGURE 11
The human behavior patterns in a fog environment, which demonstrate wall following **(A)**, area coverage **(B)**, strategy switching **(C)**, and random walk **(D)** behaviors.

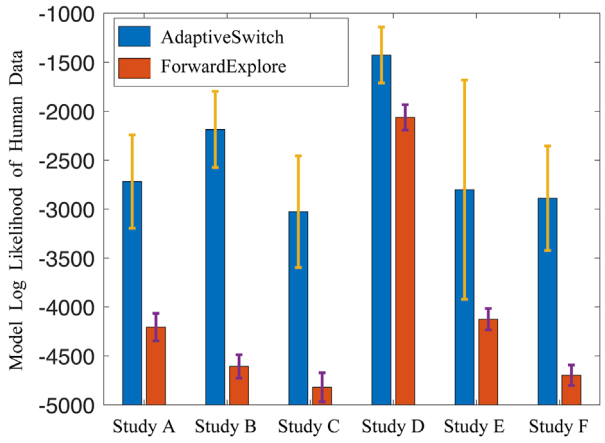


FIGURE 12
Averaged model log likelihood of AdaptiveSwitch and ForwardExplore in six human studies.

deteriorates as time pressure increases. Under moderate TP, the three time-adaptive strategies use fewer features but achieve better classification performance than Bayes optimal. This finding exemplifies the less-can-be-more effect (Gigerenzer and Gaissmaier, 2011). The classification efficiency measures the average contribution of each feature to the classification performance. Bayes optimal displays the lowest efficiency, because it utilizes all features for all time pressure conditions, whereas H_{LogOdds} exhibits the highest efficiency among the three heuristics across all time pressure conditions.

7 Mathematical modeling of human active satisficing strategies

In the active satisficing experiments, human participants face pressures due to (unmodelled) information cost (money) and sensory deprivation (fog pressure). These pressures prevent the participants from performing the test and action decisions optimally. The data analysis results for the information cost pressure, as described in Section 7.1, reveal that the test decisions and action decisions are coupled. The pressure on test decisions affect the action decisions made by the participants. The data analysis of the sensory deprivation (fog pressure) does not incorporate existing decision-making models, such as Ziebart et al. (2008); Levine et al. (2011); Ghahramani (2006); Puterman (1990), because the human participants perceive very limited information, thus violating the assumptions underlying these models. Instead, a set of decision rules are extracted in the form of heuristics from the human participants data from inspection. These heuristics capture the decision-making strategies used by the participants under sensory deprivation (fog pressure).

7.1 Information cost (money) pressure

Previous studies showed that, when information cost was present, humans used a single good reason strategy (e.g., take-the-best) in larger proportion than compensatory strategies, which integrated all available features, to make decisions (Dieckmann and Rieskamp, 2007); and information cost induced humans to optimize decision criteria and shift strategies to save cost on inferior features (Bröder, 2003). This section analyzes the characteristics of human decision behavior under information cost pressure compared with no pressure condition.

Based on the classic "treasure hunt" problem formulation for active perception (Ferrari and Wettergren, 2021), the goals of action and test decisions are expressed through three objectives, namely, information value or benefit (B), information cost (J), and distance travelled (D). Hence, optimal strategies are typically assumed to maximize a weighted sum of the three objectives, i.e.,

$$V = \sum_{k=0}^{T} \omega_B B(t_k) - \omega_D D(t_k) - \omega_J J(t_k) \tag{8}$$

where, the weights ω_B, ω_D, and ω_J represent the relative importance of the corresponding objectives.

Upon entering the study, human participants are instructed to solve the treasure hunt problem by maximizing the number of treasures found using minimum time (distance) and money. Therefore, it can be assumed that human participants also seek to maximize the objective function in (23), using their personal criteria for relative importance and decision strategy. Since the mathematical form of the chosen objectives is unknown, upon trial completion the averaged weights utilized by human participants are estimated using the Maximum Entropy Inverse Reinforcement Learning algorithm, adopted from Ziebart et al. (2008). The learned weights can then be used to understand the effects of money pressure on human decision behaviors, as follows. The two indices, $I_{IG} = \omega_B/\omega_D$ and $I_{IC} = \omega_B/\omega_J$, are obtained from the ratios of the three averaged weights and, thus, reflect the priorities underlying human decisions and behaviors. The first index, I_{IG}, referred to as information-value attempt index, measures the willingness of human participants to trade travel distance in favor of increased information value. The second index, I_{IC}, referred to as information-cost parsimony index, measures the willingness of human participants to spend "money" in favor of increased information value.

The analysis of human experiment data, shown in Figure 9, indicates that, under information cost (money) pressure, human participants are willing to travel longer distances to acquire information of high value ($\uparrow I_{IG}$). However, they are less willing to incur costs ($\downarrow I_{IC}$) for information value, thus suggesting a tendency to be more conservative in spending resources for information acquisition. Furthermore, assuming no other utility (goal) is associated with human states or actions, the causal relationships underlying human decisions may be modeled using dynamic Bayesian networks (DBNs) learned from the human trials. The DBN intra-slice structure, shown in Figure 10, uses nodes to represent the human participants' states $\mathbf{q}_k$, action decision $a(t_k)$, test decision $u(t_k)$, the set of visible targets $o(t_k)$ at time t_k, and the "money"(information cost) already spent $J(t_k)$. The intra-slice variables capture the relevant information for decision-making at a specific time slice, learning both arcs and parameters from human data.

Once the DBN description of human decisions is obtained, the inter-slice structure may be used to understand how observations influence subsequent action and test decisions. The

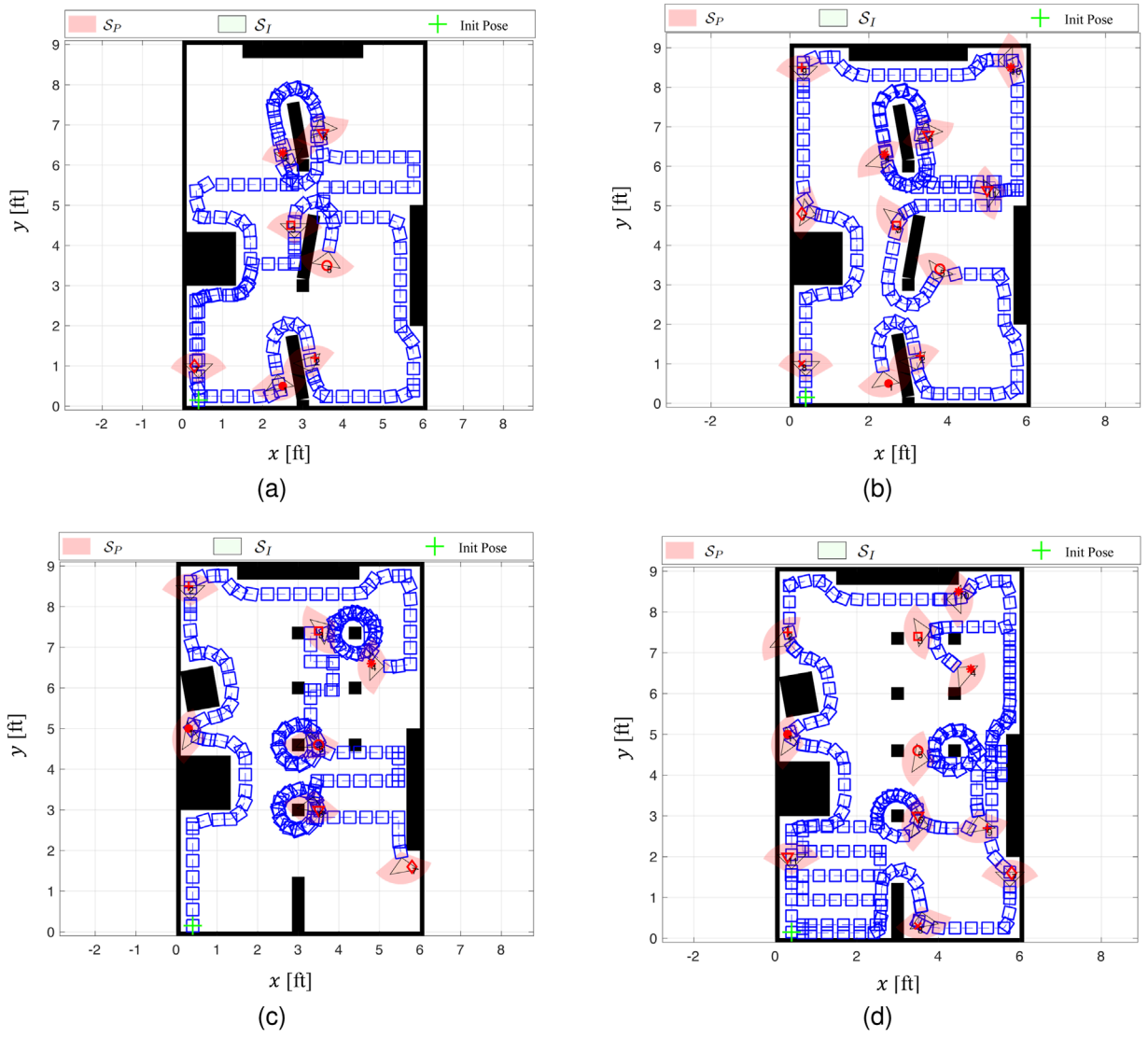


FIGURE 13
Four workspace in MATLAB® simulations and AdaptiveSwitch trajectories for case studies (*a*) – (*d*): with 7 targets plus 9 obstacles **(A)**, 11 targets plus 9 obstacles **(B)**, 7 targets plus 12 obstacles **(C)**, and 11 targets plus 12 obstacles **(D)**.

key question is: in how many time slices does an observation $o(t_k)$ influence decision-making? To determine the appropriate inter-slice structure, this paper conducts a series of hypothesis tests to assess the conformity of various models against the human decision data. Supplementary Figure S3 presents the results of these hypothesis tests. Each data point represents a p-value that evaluates the null hypothesis: "model $i+1$ does not fit the human data significantly better than model i". The models are defined according to the number of time slices in which an observation influences decisions. If the p-value is smaller than the significance level α, the null hypothesis is rejected, thus indicating that the subsequent model fits the data better than the previous one.

According to the results plotted in Supplementary Figure S3, under the no pressure condition, an observation $o(t_k)$ influences one subsequent decision. However, under the information cost(money) pressure, an observation $o(t_k)$ influences nine subsequent decisions. This finding suggests that the influence of observations extends over a longer time horizon under information cost(money) pressure than in the no pressure condition.

7.2 Sensory deprivation (fog pressure)

The introduction of sensory deprivation (fog pressure) in the environment poses two main difficulties for human participants during navigation. First, fog limits the visibility range, thus hindering human participants' capability of locating targets and being aware of obstacles. Second, fog impairs spatial awareness, thus hindering human participants' ability to accurately perceive their own position within the workspace.

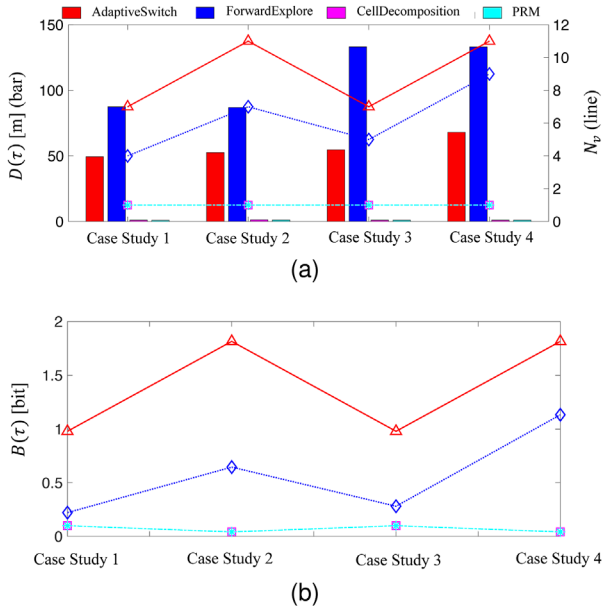


FIGURE 14
(A) Number classified targets and travel distance **(B)**information gain for two heuristic strategies and two existing robot strategies in four case studies.

In situations in which target and obstacle information is scarcely accessible and uncertainties are difficult to model, human participants were found to use local information to navigate the workspace, observe features, and classify all targets in their FOVs (Gigerenzer and Gaissmaier, 2011; Dieckmann and Rieskamp, 2007). By analyzing the human decision data collected through the active satisficing experiments described in Section 3, significant behavioral patterns shared by the top human performers can be summarized by the following six behavioral patterns exemplified by the sample studies plotted in Figure 11:

1. When participants enter an area and no targets are immediately visible, they follow the walls or obstacles detected in the workspace (Figure 11A).
2. When participants detect multiple targets, they pursue targets one by one, prioritizing them by proximity (Figure 11B).
3. While following a wall or obstacle, if participants detect a target, they will deviate from their original path and pursue the target, and may then return to their previous "wall/obstacle follow" path after performing classification (Figure 11A).
4. Upon entering an enclosed area (e.g., room), participants may engage in a strategy of covering the entire room (Figure 11B).
5. After walking along a wall or obstacle for some time without encountering any targets, participants are likely to switch to a different exploratory strategy (Figure 11C).
6. In the absence of any visible targets, participants may exhibit random walking behavior (Figure 11D).

Detailed analysis of the above behavioral patterns (omitted for brevity) showed that the following three underlying incentives drive human participants in the presence of fog pressure:

- Frugal: Human participants exhibit tendencies to avoid repeated visitations. Navigating along walls or obstacles helps participants localize themselves by using walls or obstacles as reference points.
- Greedy: Human participants demonstrate a strong motivation to find targets and engage with them. After a target is detected, participants pursue it and interact with it immediately.
- Adaptive: Human participants display adaptability by using multiple strategies for exploring the workspace. These strategies include "wall/obstacle following," "area coverage," and "random walk." Participants can switch among these strategies according to the effectiveness of their current approach in finding targets.

Based on these findings, a new algorithm referred to as AdaptiveSwitch (Algorithm 1) was developed to emulate humans' ability to transition between the three heuristics when sensory deprivation prevents the implementation of optimizing strategies. The three exploratory heuristics consist of wall/obstacle following (π_1), area coverage (π_2), and random walk (π_3). The probability of executing each heuristic is referred to as $\Pi = [b_1, b_2, b_3]^T$, where b_i represents the probability of executing π_i. The index g indicates the exploratory policy being executed, and k represents the number of steps taken while executing a policy. The maximum number of steps before updating the distribution Π is K. The policy for interacting with targets is $\pi_I(u(t_k) \mid \mathbf{q}_k, o(t_k))$, and the policy for pursuing a target is $\pi_P(a(t_k) \mid \mathbf{q}_k, o(t_k))$.

As shown in Algorithm 1, the greediness of the heuristic strategy (lines 4–9) captures the behaviors in which participants interact with targets if possible (line 4) and pursue a target if it is visible (line 7). If no targets are visible and the maximum exploratory step K is not exceeded, the current exploratory heuristic continues to be executed (lines 11–13). The adaptiveness of the three exploratory heuristics is shown in lines 15–22. If the current exploratory heuristic is executed for more than K steps, its probability of execution is discounted (line 16). The probability of executing the "wall/obstacle following" heuristic increases $\beta > 1.0$ if the participant is close to a wall/obstacle; otherwise this heuristic is disabled (lines 19–21).

After learning the parameters from the human data, the AdaptiveSwitch algorithm was compared to another hypothesized switching logic referred to as ForwardExplore in which participants predominantly move forward with a high probability and turn with a small probability or when encountering an obstacle. In order to determine which switching logic best captured human behaviors, the log likelihood of AdaptiveSwitch and ForwardExplore was computed using the human data from the active satisficing experiment involving six participants. The results plotted in Figure 12 show that the log likelihood of AdaptiveSwitch is greater than that of ForwardExplore across all human experiment trials. This finding suggests that AdaptiveSwitch aligns more closely with the observed human strategies than ForwardExplore and, therefore, was implemented in the robot studies described in the next section.

TABLE 2 Performance comparison of AdaptiveSwitch and Standalone heuristics in Webots® : Workspace A.

Performance metrics	Heuristic strategies		
	AdaptiveSwitch	RandomWalk	AreaCoverage
Travel distance, $D(\tau)$ [m]	86.19	164.87	224.18
Number of classified targets, N_v	7/7	7/7	3/7
Target visitation efficiency, η_v [m^{-1}]	0.0812	0.0425	0.0134
Travel distance, $D(\tau)$ [m]	148.98	291.69	246.38
Number of classified targets, N_v	13/13	11/13	6/13
Target visitation efficiency, η_v [m^{-1}]	0.0873	0.0377	0.0244
Travel distance, $D(\tau)$ [m]	159.97	236.86	205.78
Number of classified targets, N_v	15/15	11/15	8/15
Target visitation efficiency, η_v [m^{-1}]	0.0938	0.0464	0.0389

TABLE 3 Performance comparison of AdaptiveSwitch and Standalone heuristics in Webots® : Workspace B.

Performance metrics	Heuristic strategies		
	AdaptiveSwitch	RandomWalk	AreaCoverage
Travel distance, $D(\tau)$ [m]	122.86	218.72	265.49
Number of classified targets, N_v	7/7	5/7	5/7
Target visitation efficiency, η_v [m^{-1}]	0.0570	0.0229	0.0188
Travel distance, $D(\tau)$ [m]	122.57	219.49	234.70
Number of classified targets, N_v	13/13	10/13	7/13
Target visitation efficiency, η_v [m^{-1}]	0.0873	0.0456	0.0298
Travel distance: $D(\tau)$ [m]	129.19	226.57	216.25
Number of classified targets, N_v	15/15	12/15	8/15
Target visitation efficiency, η_v [m^{-1}]	0.1161	0.0530	0.0370

8 Autonomous robot applications of active satisficing strategies

Two key contributions of this paper are the applications of the modeled human strategies on a robot, and the comparison of optimal strategies and the modeled human strategies in pressure conditions, under which optimization is infeasible. For simplicity, the preferred sensing directions of $\mathcal{S}_P$ and $\mathcal{S}_I$ are assumed to be fixed with respect to the robot platform. Therefore, the state vector for a robot reduces to $\mathbf{q} = [x \quad y \quad \theta]^T$, where the orientation of the robot platform θ also represents the preferred sensing directions. Both sensor FOVs are modeled by sectors with angle-of-view $\zeta_1, \zeta_2 \in [0, 2\pi)$ and radii $r_1, r_2 > 0$. The two FOVs share the same apex and their bisectors coincide with each other.

8.1 Information cost (money) pressure

The introduction of information cost increases the complexity of planning test decisions. In the absence of information cost, a greedy policy that observes all available features for any target is considered "optimal", because it collects all information value without any cost. However, when information cost is taken into account, a longer planning horizon for test decisions becomes crucial to effectively allocate the budget for observing features of all targets. This paper implements two existing robot planners, PRM and cell decomposition, to solve the treasure hunt problem in an identical workspace, initial conditions, and target layouts faced by human participants in the active satisficing treasure hunt experiment. The objective function Equation 8 is maximized by

FIGURE 15
Object detection results **(A)** in clear and **(B)** fog conditions.

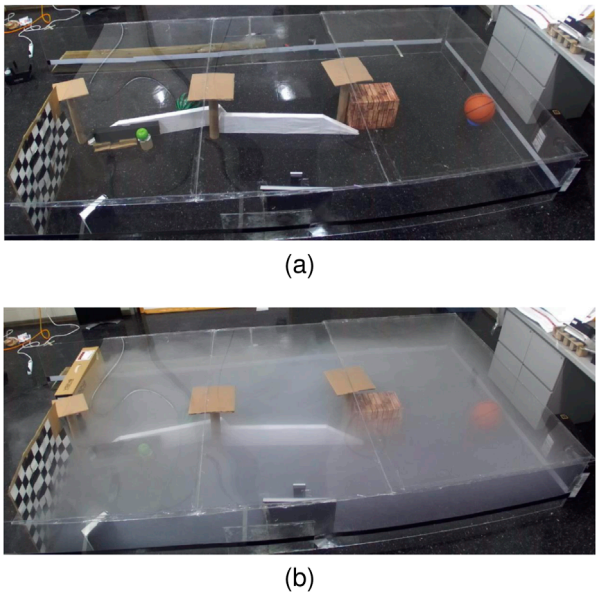


FIGURE 16
The first workspace and target layout for the physical experiment under **(A)** clear and **(B)** fog condition.

using these methods. Unlike existing approaches (Ferrari and Cai, 2009; Cai and Ferrari, 2009; Zhang et al., 2009) that solve the original version of the treasure hunt problem as described in (Ferrari and Wettergren, 2021), the developed planners handle the problem without pre-specification of the final robot configuration. Consequently, the search space increases exponentially, thus rendering label-correcting algorithms (Bertsekas, 2012) no longer applicable. Additionally, unlike previous methods that solely optimize the objective with respect to the path, the developed planners consider the constraint on the number of observed features due to information cost pressure. The number of observed features thus becomes a decision variable with a long planning horizon. To solve the problem, the developed planners use PRM and cell decomposition techniques to generate graphs representing the workspace (Ferrari and Wettergren, 2021). The Dijkstra algorithm is used to compute the shortest path between targets. Furthermore, an MINLP algorithm is used to determine the optimal number of observed features and the visitation sequence of the targets.

8.1.1 Performance comparison with human strategies

The performance of the optimal strategies known as PRM and cell decomposition is compared to that of human strategies in Supplementary Figure S4. It can be seen that, under information cost (money) pressure, the path and number of observed features per target are optimized using a linear combination of three objectives. Letting τ denote the planned path (as defined in (LaValle, 2006)), four performance metrics are used for evaluation and comparison, i.e.,: path efficiency $\eta_P = 1/D(\tau)$ $[m^{-1}]$; information gathering efficiency $\eta_B = B(\tau)/D(\tau)$ [bit/m]; measurement productivity $\eta_J = B(\tau)/J(\tau)$ [bit]; and classification performance $N = N(\tau)$ (with higher values indicating higher performance). Six case studies are examined. One case study comprises of three different experiment layouts. The optimal strategies and the human participants have no prior knowledge of the target positions and initial features, and all environmental information is obtained from FOV $\mathcal{S}_p$. The results, shown in Supplementary Figure S4, indicate that the two optimal strategies consistently outperform the human strategy across all four performance metrics. The performance envelopes of the optimal strategies are outside of the performance of the human strategy, thus indicating their superiority.

The finding that the optimal strategies outperform human strategies is unsurprising, because information cost (money) pressure imposes a constraint on only the expenditure of measurement resources, which can be effectively modeled mathematically. The finding suggests that under information cost (money) pressure, near-optimal strategies can make better decisions than human strategies.

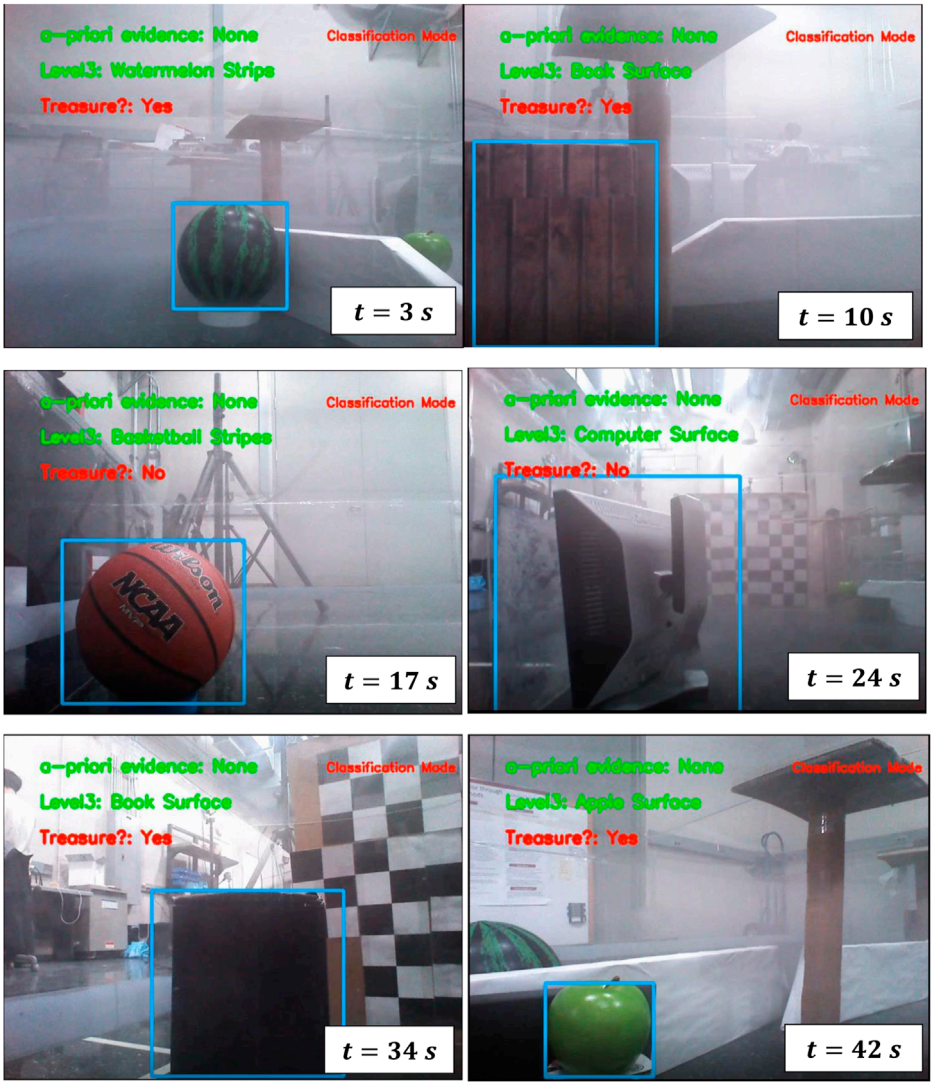


FIGURE 17
Target visitation sequence of AdaptiveSwitch in the first workspace.

TABLE 4 Performance Comparison of Heuristic Strategies in target layout 1.

Performance metrics	Heuristic strategies	
	AdaptiveSwitch	ForwardExplore
Number of classified targets, N_v	6/6	6/6
Travel distance, $D(\tau)$ [m]	6.43 ± 0.90	8.38 ± 2.07
Correct target feature classifications	13.40 ± 1.82	12.40 ± 1.95
Info gathering efficiency, η_B [bit/m]	0.155 ± 0.023	0.090 ± 0.018

8.2 Sensory deprivation (fog pressure)

An extensive series of tests are conducted to evaluate the effectiveness of AdaptiveSwitch (Section 7.) under sensory deprivation(fog) conditions and compare it with other strategies. These tests comprise of 118 simulations and physical experiments, encompassing various levels of uncertainty. The challenges posed by fog in robot planning are twofold. First, fog obstructs the robot's ability to detect targets and obstacles by using onboard sensors such as cameras, thus making long-horizon optimization-based planning nearly impossible. Second, fog complicates the task of self-localization for the robot with respect to the entire map, although short-term localization can rely on inertial measurement units. Three test groups are described as follows:

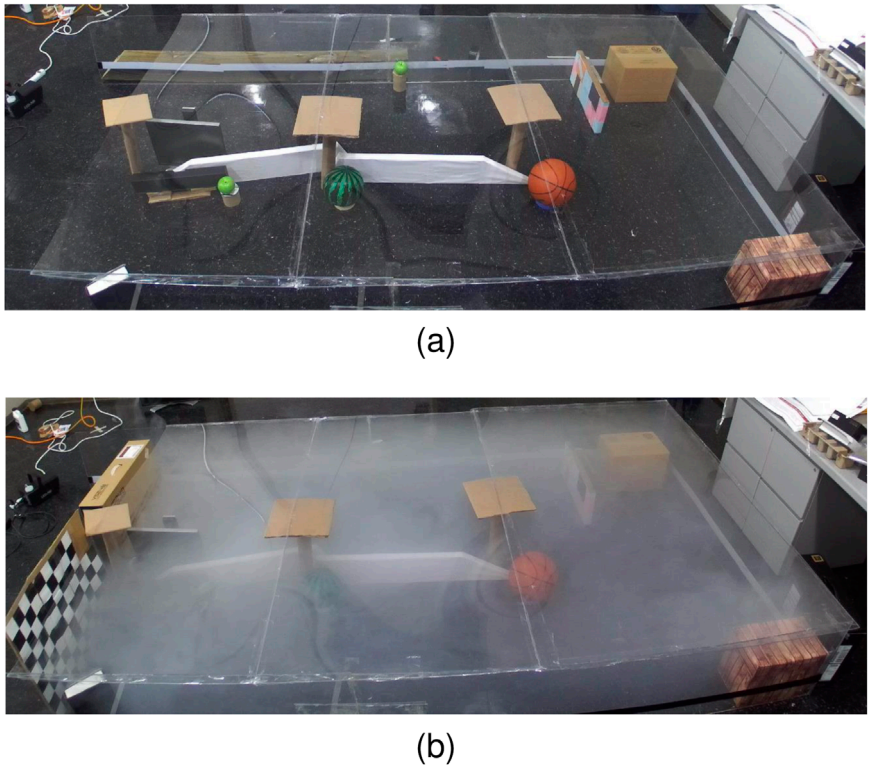


FIGURE 18
The second workspace and target layout for the physical experiment under **(A)** clear and **(B)** fog condition.

8.2.1 Performance comparison tests inside human experiment workspace

AdaptiveSwitch is applied to robots operating in the same workspace and target layouts used in the active satisficing human experiments (Section 3.), described in Figures 3, 4. Using these eighteen environments, the performance of hypothesized human strategies, AdaptiveSwitch and ForwardExplore, was compared to that of existing robot strategies (cell decomposition and PRM). One important metric used to evaluate a strategy's capability to search for targets in fog conditions is the number of classified targets: N_v. As shown by the quantitative comparison in Supplementary Figure S5A and Supplementary Figure S5B, under sensory deprivation (fog) the optimal strategies face difficulties in moving and classifying targets because of the lack of information on target and obstacle layout. Here, the distance travelled and classification performance are plotted by averaging the results of extensive simulations, along with the standard deviation (bars in Supplementary Figures S5A, S5B). In contrast, both the human strategies and AdaptiveSwitch are able to explore the unknown environment, even if at times they do not capture target information through $\mathcal{S}_p$. In particular, AdaptiveSwitch achieves slightly higher target classification rates and shorter travel distances than the observed human strategies.

8.2.2 Generalized performance comparison

In order to demonstrate the generalizability of the human-inspired strategy AdaptiveSwitch to robot applications, extensive comparative studies were performed using new workspaces and target layouts, different from those used in human experiments. In order to fully assess the performance and generalizability of AdaptiveSwitch, the sensor range was also varied to investigate the influence of sensor modalities and characteristics. Extensive simulations were conducted in MATLAB® using four newly designed workspaces and corresponding target layouts (Figure 13). For evaluation purposes, these additional simulations considered fixed FOV geometries and assumed no missed detections or false alarms, as well as perfect target feature recognition.

As part of this comparison, ForwardExplore and the two existing robot strategies, cell decomposition and PRM, are also implemented for comparison. Due to the limitations posed by fog and limited sensing capabilities, the performance in terms of travel distance, $D(\tau)$, and classification, (N_v), is significantly hindered, as shown by the averaged values plotted in Figure 14A on the left (histogram bar) and right (line) vertical axis, respectively. Robots implementing AdaptiveSwitch outperform those implementing other strategies in terms of the number of correctly classified targets, because they are able to explore the workspace even when no targets were visible.

Additionally, AdaptiveSwitch is more efficient than ForwardExplore in terms of travel distance. By adapting its exploration strategy and leveraging the combination of three simple heuristics, AdaptiveSwitch is able to classify more targets while traveling shorter distances. Consequently, higher information value $B(\tau)$ than that with both ForwardExplore and existing robot strategies is observed across all four case studies (Figure 14B). These findings highlight the effectiveness of the AdaptiveSwitch in navigating foggy environments and its superiority to existing

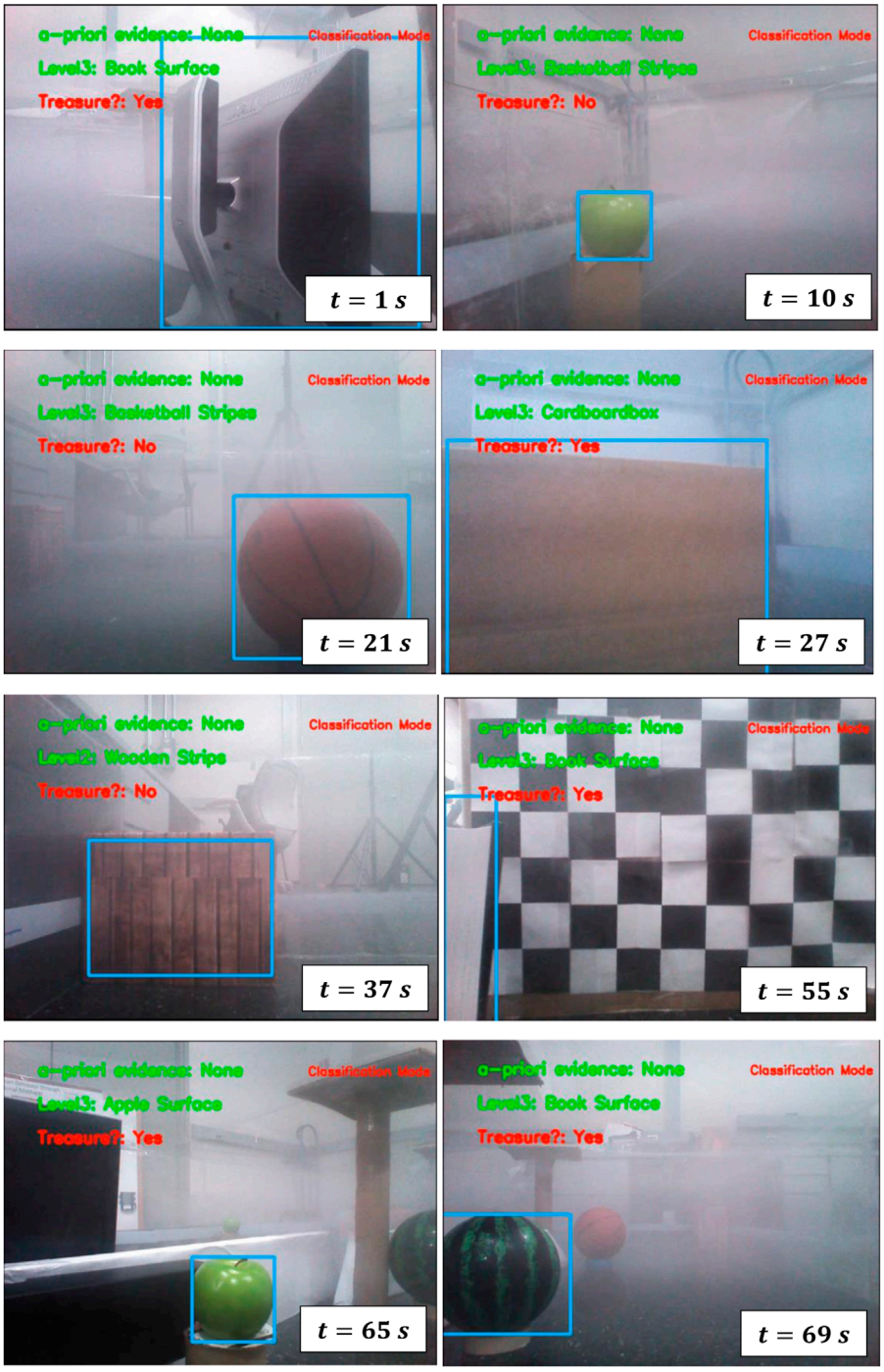


FIGURE 19
Target visitation sequence of AdaptiveSwitch in the second workspace.

robot strategies and the ForwardExplore in terms of information gathering and travel efficiency.

8.2.2.1 Simulations with artificial fog

Two new workspaces are designed in Webots® as shown in Supplementary Figure S6. The performance of AdaptiveSwitch and its standalone heuristics for the two workspaces is shown in Tables 2, 3. The comparison reveals the substantial advantage of AdaptiveSwitch. In both workspace scenarios, as shown in Tables 2, 3, AdaptiveSwitch outperforms its standalone heuristics by successfully finding and classifying all targets within the given simulation time upper bound. In contrast, the standalone heuristics are unable to achieve this level of performance. AdaptiveSwitch not only visits and classifies all targets, but also accomplishes the tasks within shorter travel distances than the standalone heuristics. Therefore, AdaptiveSwitch exhibits higher target visitation efficiency (η_v) which is calculated as the ratio of the

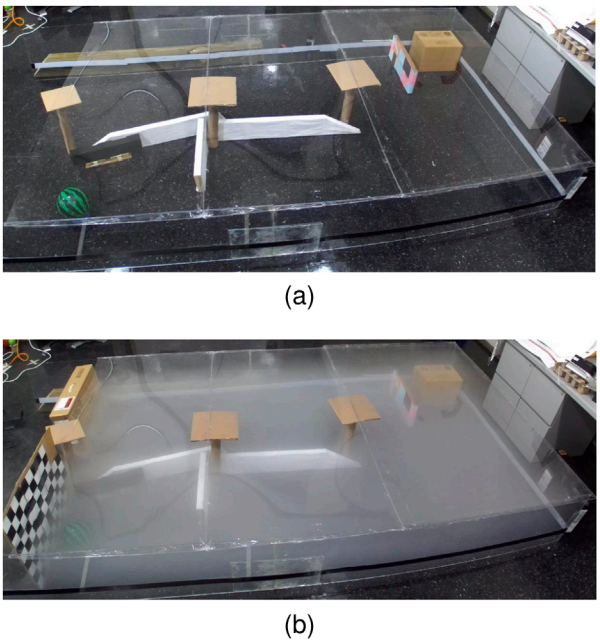


FIGURE 20
The third workspace and target layout for the physical experiment under **(A)** clear and **(B)** fog condition.

number of classified targets to the travel distance ($N_v/D(\tau)$). The target visitation efficiency of AdaptiveSwitch is at least twice higher than that of the standalone heuristics thanks to the combination of multiple, simple heuristics. In contrast, when used in a stand-alone fashion, the same heuristics may become trapped in ineffective "moving patterns", struggling to perform in certain areas of the workspace.

8.2.3 Physical experiment tests in real fog environment

To handle real-world uncertainties that are not adequately modeled in simulations, this paper conducts physical experiments to test the AdaptiveSwitch. These uncertainties include factors such as the robot's initial position and orientation, target miss detection and false alarms, depth measurement errors, and control disturbances. In addition, the fog models available in Webots®, are relatively simple and do not provide a wide range of possibilities for simulating the degrading effects of fog on target detection and classification performance. Consequently, this paper performs physical experiments to better capture the complexities and uncertainties associated with real-world conditions.

The physical experiments use the ROSbot2.0 robot equipped with an RGB-D camera as the primary sensor. The YOLOv3 object detection algorithm, the best applicable at the time of these studies, was implemented to detect targets of interest (e.g., an apple, watermelon, orange, basketball, computer, book, cardboard box, and wooden box) identical to those in human experiments. Training images for the YOLOv3 were obtained in fog-free environments, in order to later test the robot's ability to cope with unseen conditions (fog pressure) in real time.

As shown in Figure 15, the YOLOv3 algorithm successfully detects the existence of the target "computer" when the environment is clear, as shown in Figure 15A. However, when fog is present, as illustrated in Figure 15B, the algorithm fails to detect the target. This result demonstrates the degrading effect on the performance of target detection algorithms.

In the physical experiments conducted with ROSbot2.0 (Husarion, 2018), AdaptiveSwitch and ForwardExplore are implemented to test their performance in an environment with fog. A plastic box is constructed with dimensions 10′0″ x 6′0″ x 1′8″ in order to create the foggy environment. The box is designed to contain different layouts of obstacles and targets, capturing various aspects of a "treasure hunt" scenario, such as target density and target view angles. Each heuristic strategy is tested five times in each layout, considering all the uncertainties described earlier. The travel distances in the physical experiments are measured in inertial measurement unit.

The first layout (Figure 16) is comprised of six targets, i.e.,: a watermelon, wooden box, basketball, book, apple, and computer. The target visitation sequences of AdaptiveSwitch along the path are depicted in Figure 17, showing the robot's trajectory and the order in which the targets are visited. The performance of the two strategies is summarized in Table 4, as evaluated according to three aspects: travel distance $D(\tau)$, correct target feature classifications, and information gathering efficiency η_B. These metrics assess the quality of the strategies' action and test decisions.

The second layout (Figure 18) contains eight targets: a watermelon, wooden box, basketball, book, computer, cardboard box, and two apples. The obstacles layout is also changed with respect to the first layout: the cardboard box is placed in a "corner" and is visible from only one direction, thus increasing the difficulty of detecting this target. This layout enables a case study in which the targets are more crowded than in the first layout. The mobile robot first-person-views of AdaptiveSwitch along the path are demonstrated in Figure 19, and the performance is shown in Supplementary Table S1.

The third layout (Figure 20) contains two targets: a cardboard box, and a watermelon. Note that having fewer targets does not necessarily make the problem easier, because the difficulty in target search in fog comes from how to navigate when no target is in the FOV. This layout intentionally makes the problem "difficult", because it "hides" two targets behind the walls. The mobile robot first-person-views of AdaptiveSwitch along the path are demonstrated in Figure 21, and the performance is shown in Supplementary Table S2. The videos for all physical experiments (AdaptiveSwitch and ForwardExplore in three layouts) are accessible through the link in (Chen, 2021).

According to the performance summaries in Table 4, Supplementary Tables S1, S2, both AdaptiveSwitch and ForwardExplore are capable of visiting and classifying all targets in the three layouts under real-world uncertainties. However, AdaptiveSwitch demonstrates several advantages over ForwardExplore:

1. The average travel distance of AdaptiveSwitch is 30.33%, 59.93%, and 56.02% more efficient than ForwardExplore in the three workspaces, respectively. This finding indicates that AdaptiveSwitch is able to search target with a shorter travel distance than ForwardExplore.

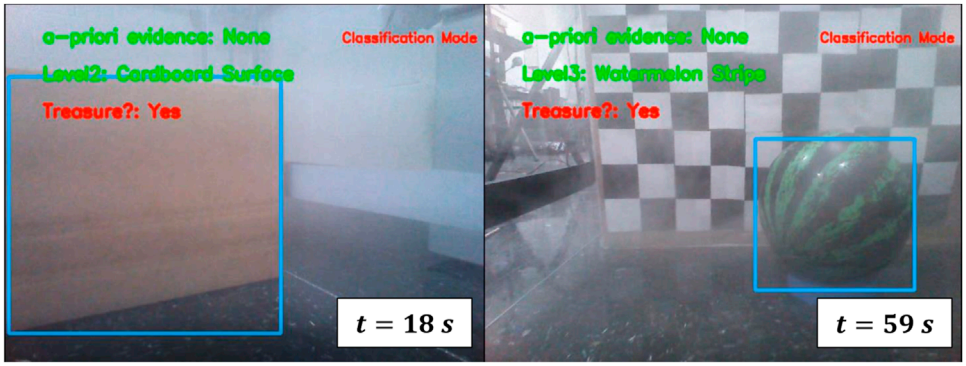


FIGURE 21
Target visitation sequence of AdaptiveSwitch in the third workspace.

$\Pi = [b_1, b_2, b_3]^T$
$k = 0, \quad g = 0$
while ($t_k \leq t_T \vee$ not all targets are classified) **do**
  if $\exists \mathbf{x}_j \in \mathcal{S}_I(\mathbf{q}_k)$ **then**
    $\pi_I(u(t_k) \mid \mathbf{q}_k, o(t_k))$
  else
    if $o(t_k) \neq \emptyset$ **then**
      $\pi_P(a(t_k) \mid \mathbf{q}_k, o(t_k))$
      $k = 0, \quad g = 0$
    else
      if $g > 0 \wedge k \leq K$ **then**
        $\pi_g(a(t_k) \mid \mathbf{q}_k, o(t_k))$
        $k = k + 1$
      else
        if $k \geq K$ **then**
          $\Pi[g] = \gamma * \Pi[g]$
        else
          if not closed to wall **then**
            $\Pi[1] = 0$
          else
            $\Pi[1] = \beta(b_1 + b_2 + b_3)$
          end if
        end if
        $\Pi = \text{normalize}(\Pi)$
        $g \sim \Pi$
      end if
    end if
  end if
end while

Algorithm 1. AdaptiveSwitch.

2. The target feature classification performance of AdaptiveSwitch is slightly better than that of ForwardExplore, with improvements of 8.06%, 17.11%, and 4.16% in the three workspace, respectively. One possible explanation for these results is that the "obstacle follow" and "area coverage" heuristics in AdaptiveSwitch cause the robot's body to be parallel to obstacles during classification of target features, thus ensuring that the targets are the major part of the robot's first-person view and make them relatively easier to classify. In contrast, ForwardExplore does not always lead the robot body to be parallel to obstacles during classification, thereby sometimes allowing obstacles to dominate the robot's first-person view and decreasing the target classification performance.

9 Summary and conclusion

This paper presents novel satisficing solutions that modulate between near-optimal and heuristics to solve satisficing treasure hunt problem under environment pressures. These proposed solutions are derived from human decision data collected through both passive and active satisficing experiments. The ultimate goal is to apply these satisficing solutions to autonomous robots. The modeled passive satisficing strategies adaptively select target features to be entered in measurement model based on a given time pressure. The idea behind this approach is the human participants behavior that dropping less informative features for inference in order to meet the decision deadline. The results show that the modeled passive satisficing strategies outperform the "optimal" strategy that always use all available features for inference in terms of classification performance and significantly reduce the complexity of target feature search compared with exhaustive search.

Regarding the active satisficing strategies, the strategy that deals with information cost formulates an optimization problem with the hard constraint imposed by information cost. This approach is taken because the information cost constraint doesn't fundamentally undermine the accuracy of the model of the world and the agent, and optimization still yield high-quality decisions. The results show that the strategy outperforms human participants across several key metrics (e.g., travel distance and measurement productivity, etc.). However, under sensory deprivation, the knowledge of the world is severely compromised, and thus decisions produced by optimization is risky or even no longer feasible, which is also demonstrated

through experiments in this paper. The modeled human strategies named AdaptiveSwitch shows the ability to use local information and navigate in foggy environment by using heuristics derived from humans. The results also show that the AdaptiveSwitch can adapt to varying workspaces with different obstacle layouts, target density, etc., beyond the workspace used in the active satisficing experiments. Finally, AdaptiveSwitch is implemented on a physical robot and conducts satisificing treasure hunt with actual fog, which demonstrates the ability to deal with real-life uncertainties in both perception and action.

Overall, the proposed satisficing strategies comprise of a toolbox, which can be readily deployed on a robot in order to address different real-life environment pressures encountered during the mission. These strategies provide solutions to scenarios characterized by time limitations, constraints on available resources (e.g., fuel or energy), and adverse weathers such as fog or heavy rain.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YC: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review and editing. PZ: Data curation, Investigation, Methodology, Software, Validation, Writing–review and editing. AA: Investigation, Software, Writing–review and editing. TE: Investigation, Software, Writing–review and editing. MS: Investigation, Software, Writing–review and editing. SF: Funding acquisition, Project administration, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research is funded by the Office of Naval Research (ONR) Science of Autonomy Program, under Grant N00014-13-1-0561.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2024.1384609/full#supplementary-material

References

Abdulsaheb, J. A., and Kadhim, D. J. (2023). Classical and heuristic approaches for mobile robot path planning: a survey. *Robotics* 12 (4), 93. doi:10.3390/robotics12040093

Batta, E., and Stephens, C. (2019). "Heuristics as decision-making habits of autonomous sensorimotor agents," in Artificial Life Conference Proceedings. Cambridge, MA, USA: MIT Press, 72–78.

Bertsekas, D. (2012). *Dynamic programming and optimal control: volume I*, 1. Belmont, MA: Athena scientific.

Bishop, C. M., and Nasrabadi, N. M. (2006). Pattern recognition and machine learning. *Springer* 4 (4). doi:10.1007/978-0-387-45528-0

Brighton, H., and Gigerenzer, G. (2008). "Bayesian brains and cognitive mechanisms: harmony or dissonance," in *The probabilistic mind: prospects for Bayesian cognitive science*. Editors N. Chater, and M. Oaksford, 189–208.

Bröder, A. (2003). "Decision making with the" adaptive toolbox": influence of environmental structure, intelligence, and working memory load. *J. Exp. Psychol. Learn. Mem. Cognition* 29 (4), 611–625. doi:10.1037/0278-7393.29.4.611

Cai, C., and Ferrari, S. (2009). Information-driven sensor path planning by approximate cell decomposition. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* 39 (3), 672–689. doi:10.1109/tsmcb.2008.2008561

Caplin, A., and Glimcher, P. W. (2014). "Basic methods from neoclassical economics," in *Neuroeconomics* (Elsevier), 3–17.

Chen, D., Zhou, B., Koltun, V., and Krähenbühl, P. (2020). "Learning by cheating," in *Conference on robot learning* (Cambridge, MA: PMLR), 66–75.

Chen, Y. (2021). *Navigation in fog*. Cornell University. Available at: https://youtu.be/b9Cca0XSxAQ.

Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29 (37), 11560–11571. doi:10.1523/jneurosci.1844-09.2009

Dieckmann, A., and Rieskamp, J. (2007). The influence of information redundancy on probabilistic inferences. *Mem. and Cognition* 35 (7), 1801–1813. doi:10.3758/bf03193511

Dua, D., and Graff, C. (2017). UCI machine learning repository. Available at: http://archive.ics.uci.edu/ml.

Ferrari, S., and Cai, C. (2009). Information-driven search strategies in the board game of CLUE. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* 39 (3), 607–625. doi:10.1109/TSMCB.2008.2007629

Ferrari, S., and Vaghi, A. (2006). Demining sensor modeling and feature-level fusion by bayesian networks. *IEEE Sensors J.* 6 (2), 471–483. doi:10.1109/jsen.2006.870162

Ferrari, S., and Wettergren, T. A. (2021). *Information-driven planning and control*. MIT Press.

Fishburn, P. C. (1981). Subjective expected utility: a review of normative theories. *Theory Decis.* 13 (2), 139–199. doi:10.1007/bf00134215

Garlan, D., Siewiorek, D. P., Smailagic, A., and Steenkiste, P. (2002). Project aura: toward distraction-free pervasive computing. *IEEE Pervasive Comput.* 1 (2), 22–31. doi:10.1109/mprv.2002.1012334

Ge, S. S., Zhang, Q., Abraham, A. T., and Rebsamen, B. (2011). Simultaneous path planning and topological mapping (sp2atm) for environment exploration and goal oriented navigation. *Robotics Aut. Syst.* 59 (3-4), 228–242. doi:10.1016/j.robot.2010.12.003

Gemerek, J., Fu, B., Chen, Y., Liu, Z., Zheng, M., van Wijk, D., et al. (2022). Directional sensor planning for occlusion avoidance. *IEEE Trans. Robotics* 38, 3713–3733. doi:10.1109/tro.2022.3180628

Ghahramani, Z. (2006). Learning dynamic bayesian networks. *Adapt. Process. Sequences Data Struct. Int. Summer Sch. Neural Netw. "ER Caianiello" Vietri sul Mare, Salerno, Italy Sept. 6–13, 1997 Tutor. Lect.*, 168–197. doi:10.1007/bfb0053999

Gigerenzer, G., and Brighton, H. (2009). Homo heuristicus: why biased minds make better inferences. *Top. cognitive Sci.* 1 (1), 107–143. doi:10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G., and Gaissmaier, W. (2011). Heuristic decision making. *Annu. Rev. Psychol.* 62 (1), 451–482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., and Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychol. Rev.* 103 (4), 650–669. doi:10.1037//0033-295x.103.4.650

Gigerenzer, G., and Todd, P. M. (1999). *Simple heuristics that make us smart*. USA: Oxford University Press.

Gigerenzer, G. (1991). From tools to theories: a heuristic of discovery in cognitive psychology. *Psychol. Rev.* 98 (2), 254–267. doi:10.1037//0033-295x.98.2.254

Gigerenzer, G. (2007). *Gut feelings: the intelligence of the unconscious*. Penguin.

Gluck, M. A., Shohamy, D., and Myers, C. (2002). How do people solve the "weather prediction" task? individual variability in strategies for probabilistic category learning. *Learn. and Mem.* 9 (6), 408–418. doi:10.1101/lm.45202

Goldstein, D. G., and Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychol. Rev.* 109 (1), 75–90. doi:10.1037//0033-295x.109.1.75

Herbert, A. S. (1979). "Rational decision making in business organizations," *Am. Econ. Rev.* 69 (4), 493–513.

Ho, J., and Ermon, S. (2016). Generative adversarial imitation learning. *Adv. neural Inf. Process. Syst.* 29. doi:10.5555/3157382.3157608

Hogarth, R. M., and Karelaia, N. (2007). Heuristic and linear models of judgment: matching rules and environments. *Psychol. Rev.* 114 (3), 733–758. doi:10.1037/0033-295x.114.3.733

Husarion (2018). Rosbot autonomous mobile robot. Available at: https://husarion.com/manuals/rosbot/.

Jensen, F. V., and Nielsen, T. D. (2007). *Bayesian networks and decision graphs*, 2. Springer.

Kirsch, A. (2016). "Heuristic decision-making for human-aware navigation in domestic environments," in 2nd global conference on artificial intelligence (GCAI). September 19 - October 2, 2016, Berlin, Germany.

Kruschke, J. K. (2010). Bayesian data analysis. *Wiley Interdiscip. Rev. Cognitive Sci.* 1 (5), 658–676. doi:10.1002/wcs.72

Lagnado, D. A., Newell, B. R., Kahan, S., and Shanks, D. R. (2006). Insight and strategy in multiple-cue learning. *J. Exp. Psychol. General* 135 (2), 162–183. doi:10.1037/0096-3445.135.2.162

Lamberts, K. (1995). Categorization under time pressure. *J. Exp. Psychol. General* 124 (2), 161–180. doi:10.1037//0096-3445.124.2.161

Latombe, J.-C. (2012). *Robot motion planning*, 124. Springer Science and Business Media.

LaValle, S. M. (2006). *Planning algorithms*. Cambridge, United Kingdom: Cambridge University Press.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Curr. Dir. Psychol. Sci.* 19 (3), 143–148. doi:10.1177/0963721410370295

Lebedev, M. A., Carmena, J. M., O'Doherty, J. E., Zacksenhouse, M., Henriquez, C. S., Principe, J. C., et al. (2005). Cortical ensemble adaptation to represent velocity of an artificial actuator controlled by a brain-machine interface. *J. Neurosci.* 25 (19), 4681–4693. doi:10.1523/jneurosci.4088-04.2005

Levine, S., Popovic, Z., and Koltun, V. (2011). Nonlinear inverse reinforcement learning with Gaussian processes. *Adv. neural Inf. Process. Syst.* 24. doi:10.5555/2986459.2986462

Lewis, M., and Cañamero, L. (2016). Hedonic quality or reward? a study of basic pleasure in homeostasis and decision making of a motivated autonomous robot. *Adapt. Behav.* 24 (5), 267–291. doi:10.1177/1059712316666331

Lichtman, A. J. (2008). *The keys to the White House: a surefire guide to predicting the next president*. Lanham, MD: Rowman and Littlefield.

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., et al. (2015). Continuous control with deep reinforcement learning. *arXiv Prepr. arXiv:1509.02971*. doi:10.48550/arXiv.1509.02971

Liu, C., Zhang, S., and Akbar, A. (2019). Ground feature oriented path planning for unmanned aerial vehicle mapping. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 12 (4), 1175–1187. doi:10.1109/jstars.2019.2899369

Lones, J., Lewis, M., and Cañamero, L. (2014). "Hormonal modulation of development and behaviour permits a robot to adapt to novel interactions," in Artificial Life Conference Proceedings. Cambridge, MA, USA: MIT Press, 184–191.

Martin-Rico, F., Gomez-Donoso, F., Escalona, F., Garcia-Rodriguez, J., and Cazorla, M. (2020). Semantic visual recognition in a cognitive architecture for social robots. *Integr. Computer-Aided Eng.* 27 (3), 301–316. doi:10.3233/ica-200624

Mullainathan, S., and Thaler, R. H. (2000). "Behavioral economics,".

Newell, B. R., and Shanks, D. R. (2003). "Take the best or look at the rest? factors influencing" one-reason" decision making. *J. Exp. Psychol. Learn. Mem. Cognition* 29 (1), 53–65. doi:10.1037//0278-7393.29.1.53

Nicolaides, P. (1988). Limits to the expansion of neoclassical economics. *Camb. J. Econ.* 12 (3), 313–328.

O'Brien, M. J., and Arkin, R. C. (2020). Adapting to environmental dynamics with an artificial circadian system. *Adapt. Behav.* 28 (3), 165–179. doi:10.1177/1059712319846854

Oh, H., Beck, J. M., Zhu, P., Sommer, M. A., Ferrari, S., and Egner, T. (2016). Satisficing in split-second decision making is characterized by strategic cue discounting. *J. Exp. Psychol. Learn. Mem. Cognition* 42 (12), 1937–1956. doi:10.1037/xlm0000284

Oh-Descher, H., Beck, J. M., Ferrari, S., Sommer, M. A., and Egner, T. (2017). Probabilistic inference under time pressure leads to a cortical-to-subcortical shift in decision evidence integration. *NeuroImage* 162, 138–150. doi:10.1016/j.neuroimage.2017.08.069

Pan, X., and Hamilton, A. F. d. C. (2018). Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape. *Br. J. Psychol.* 109 (3), 395–417. doi:10.1111/bjop.12290

Payne, J. W., Bettman, J. R., and Johnson, E. J. (1988). Adaptive strategy selection in decision making. *J. Exp. Psychol. Learn. Mem. Cognition* 14 (3), 534–552. doi:10.1037//0278-7393.14.3.534

Porcelli, A. J., and Delgado, M. R. (2017). Stress and decision making: effects on valuation, learning, and risk-taking. *Curr. Opin. Behav. Sci.* 14, 33–39. doi:10.1016/j.cobeha.2016.11.015

Powell, W. B. (2007). *Approximate dynamic programming: solving the curses of dimensionality*, 703. John Wiley and Sons.

Puterman, M. L. (1990). Markov decision processes. *Handb. operations Res. Manag. Sci.* 2, 331–434.

Ratcliff, R., and McKoon, G. (1989). Similarity information versus relational information: differences in the time course of retrieval. *Cogn. Psychol.* 21 (2), 139–155. doi:10.1016/0010-0285(89)90005-4

Rieskamp, J., and Otto, P. E. (2006). Ssl: a theory of how people learn to select strategies. *J. Exp. Psychol. General* 135 (2), 207–236. doi:10.1037/0096-3445.135.2.207

Rossello, N. B., Carpio, R. F., Gasparri, A., and Garone, E. (2021). Information-driven path planning for uav with limited autonomy in large-scale field monitoring. *IEEE Trans. Automation Sci. Eng.* 19 (3), 2450–2460. doi:10.1109/TASE.2021.3071251

Savage, L. J. (1972). The foundations of statistics. *Cour. Corp.*

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*.

Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.* 5 (7), 532–545. doi:10.1038/nrn1427

Servotte, J.-C., Goosse, M., Campbell, S. H., Dardenne, N., Pilote, B., Simoneau, I. L., et al. (2020). Virtual reality experience: immersion, sense of presence, and cybersickness. *Clin. Simul. Nurs.* 38, 35–43. doi:10.1016/j.ecns.2019.09.006

Si, J., Barto, A. G., Powell, W. B., and Wunsch, D. (2004). *Handbook of learning and approximate dynamic programming*, 2. John Wiley and Sons.

Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., and Riedmiller, M. (2014). "Deterministic policy gradient algorithms," in International conference on machine learning. June 21–June 26, 2014, Beijing, China. (Beijing, China: Pmlr), 387–395.

Simon, H. A., and Kadane, J. B. (1975). Optimal problem-solving search: all-or-none solutions. *Artif. Intell.* 6 (3), 235–247. doi:10.1016/0004-3702(75)90002-8

Simon, H. A. (1955). A behavioral model of rational choice. *Q. J. Econ.* 69 (1), 99–118. doi:10.2307/1884852

Simon, H. A. (1997). *Models of bounded rationality: empirically grounded economic reason*, 3. MIT press.

Simon, H. A. (2019). *The Sciences of the Artificial, reissue of the third edition with a new introduction by John Laird*. MIT press.

Slovic, P., Peters, E., Finucane, M. L., and MacGregor, D. G. (2005). Affect, risk, and decision making. *Health Psychol.* 24 (4S), S35–S40. doi:10.1037/0278-6133.24.4.s35

Speekenbrink, M., Lagnado, D. A., Wilkinson, L., Jahanshahi, M., and Shanks, D. R. (2010). Models of probabilistic category learning in Parkinson's disease: strategy use and the effects of l-dopa. *J. Math. Psychol.* 54 (1), 123–136. doi:10.1016/j.jmp.2009.07.004

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement learning: an introduction*. MIT press.

Swingler, A., and Ferrari, S. (2013). "On the duality of robot and sensor path planning," in 52nd IEEE conference on decision and control. 10-13 December 2013, Firenze, Italy. (IEEE), 984–989.

Toader, A. C., Rao, H. M., Ryoo, M., Bohlen, M. O., Cruger, J. S., Oh-Descher, H., et al. (2019). Probabilistic inferential decision-making under time pressure in rhesus macaques (macaca mulatta). *J. Comp. Psychol.* 133 (3), 380–396. doi:10.1037/com0000168

Todorov, E., and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5 (11), 1226–1235. doi:10.1038/nn963

Vallverdú, J., Talanov, M., Distefano, S., Mazzara, M., Tchitchigin, A., and Nurgaliev, I. (2016). A cognitive architecture for the implementation of emotions in computing systems. *Biol. Inspired Cogn. Archit.* 15, 34–40. doi:10.1016/j.bica.2015.11.002

Van Veen, H. A., Distler, H. K., Braun, S. J., and Bülthoff, H. H. (1998). Navigating through a virtual city: using virtual reality technology to study human action and perception. *Future Gener. Comput. Syst.* 14 (3-4), 231–242. doi:10.1016/s0167-739x(98)00027-2

Wiering, M. A., and Van Otterlo, M. (2012). Reinforcement learning. *Adapt. Learn. Optim.* 12 (3), 729.

Zhang, G., Ferrari, S., and Qian, M. (2009). An information roadmap method for robotic sensor path planning. *J. Intelligent Robotic Syst.* 56 (1), 69–98. doi:10.1007/s10846-009-9318-x

Zhang, G., Ferrari, S., and Cai, C. (2011). A comparison of information functions and search strategies for sensor planning in target classification. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* 42 (1), 2–16. doi:10.1109/TSMCB.2011.2165336

Zhu, P., Ferrari, S., Morelli, J., Linares, R., and Doerr, B. (2019). Scalable gas sensing, mapping, and path planning via decentralized hilbert maps. *Sensors* 19 (7), 1524. doi:10.3390/s19071524

Ziebart, B. D., Maas, A. L., Bagnell, J. A., and Dey, A. K. (2008). Maximum entropy inverse reinforcement learning. *Aaai* 8, 1433–1438. Chicago, IL, USA.

Zielinski, D. J., McMahan, R. P., Lu, W., and Ferrari, S. (2013). "Ml2vr: providing matlab users an easy transition to virtual reality and immersive interactivity," in 2013 IEEE virtual reality (VR). 18-20 March 2013, Lake Buena Vista, FL, USA. (IEEE), 83–84.

frontiers | Frontiers in Artificial Intelligence

TYPE Original Research
PUBLISHED 12 March 2025
DOI 10.3389/frai.2025.1540646

OPEN ACCESS

EDITED BY
Tony J. Prescott,
The University of Sheffield, United Kingdom

REVIEWED BY
Takashi Kuremoto,
Nippon Institute of Technology, Japan
Sid Ahmed Benabderrahmane,
New York University, United States

*CORRESPONDENCE
Faisal Binzagr
✉ fbinzagr@kau.edu.sa

RECEIVED 06 December 2024
ACCEPTED 11 February 2025
PUBLISHED 12 March 2025

CITATION
Binzagr F and Abulfaraj AW (2025) InGSA: integrating generalized self-attention in CNN for Alzheimer's disease classification.
Front. Artif. Intell. 8:1540646.
doi: 10.3389/frai.2025.1540646

InGSA: integrating generalized self-attention in CNN for Alzheimer's disease classification

Faisal Binzagr[1]* and Anas W. Abulfaraj[2]

[1]Department of Computer Science, King Abdulaziz University, Rabigh, Saudi Arabia, [2]Department of Information Systems, King Abdulaziz University, Rabigh, Saudi Arabia

Alzheimer's disease (AD) is an incurable neurodegenerative disorder that slowly impair the mental abilities. Early diagnosis, nevertheless, can greatly reduce the symptoms that are associated with the condition. Earlier techniques of diagnosing the AD from the MRI scans have been adopted by traditional machine learning technologies. However, such traditional methods involve depending on feature extraction that is usually complex, time-consuming, and requiring substantial effort from the medical personnel. Furthermore, these methods are usually not very specific as far as diagnosis is concerned. In general, traditional convolutional neural network (CNN) architectures have a problem with identifying AD. To this end, the developed framework consists of a new contrast enhancement approach, named haze-reduced local-global (HRLG). For multiclass AD classification, we introduce a global CNN-transformer model InGSA. The proposed InGSA is based on the InceptionV3 model which is pre-trained, and it encompasses an additional generalized self-attention (GSA) block at top of the network. This GSA module is capable of capturing the interaction not only in terms of the spatial relations within the feature space but also over the channel dimension it is capable of picking up fine detailing of the AD information while suppressing the noise. Furthermore, several GSA heads are used to exploit other dependency structures of global features as well. Our evaluation of InGSA on a two benchmark dataset, using various pre-trained networks, demonstrates the GSA's superior performance.

KEYWORDS

Alzheimer's disease classification, generalized self-attention, CNN, transfer learning, transformer

1 Introduction

Alzheimer's disease (AD) is a type of dementia that is not curable, which becomes worse over years as it affects the human brain, but early diagnosis helps to minimize the symptoms and the management of the patient (McKhann et al., 1984). Its manifestation involves impaired memory because patients cannot organize or recall information properly, and poor judgment that renders the affected persons completely helpless and in need of care as the disease develops (Choi et al., 2020). The probability raised from 2% at 65 years to 35% at 85 years for AD. Approximately 26.6 million people had it in 2006; the figure rose to over 55 million in 2020 and is expected to reach 152 million by 2050 (Gunawardena et al., 2017). Neuronal loss and synaptic impairment can occur at least one or two decades before disease onset (Böhle et al., 2019). It is essential to detect AD in the prodromal stage, which is characterized by moderate cognitive impairment (MCI), as there is currently no cure. Early MCI (EMCI) is a cognitive impairment stage that precedes MCI (Kang et al., 2020). The early detection of EMCI has the potential to

prevent the progression of EMCI to AD. The importance of diagnosing MCI patients has been emphasized by studies that has identified the distinctions between early MCI (EMCI) and late MCI (LMCI) groups (Nozadi et al., 2018; Edmonds et al., 2019; Zhang T. et al., 2019). MCI has a symptom profile that is similar, but less severe, to AD (Varatharajah et al., 2019). Nowadays, this disease is also defined as mild cognitive impairment associated with the existence of Alzheimer's disease; according to recent investigations, ∼80% of patients diagnosed with MCI develop AD in 7 years. For monitoring of variations in the densities of the brain tissues, magnetic resonance imaging (MRI) and positron emission tomography (PET) are frequently used since they do not include the invasion of the tissues (Ramzan et al., 2020; Gao, 2021). Neuroimaging, especially using MRI, is crucial for the study of the nervous system structures more closely (Tuvshinjargal and Hwang, 2022); this test helps in diagnosis of certain diseases such as tumors and cancer (Tehsin et al., 2024). MRI does work in the case of Alzheimer's; it allows capturing structural changes in the brain, for instance, the reduction of certain regions and the appearance of new formations, heterogeneous density, and the presence of abnormal substances typical of the disease (Simic et al., 2009).

In recent years, medical imagery such as MRI has been used with machine learning (ML) and deep learning (DL). These methods are used in health checks and early AD diagnosis. They also excel at categorizing images in health and computer vision (Nasir et al., 2021, 2020, 2022; Yousafzai et al., 2024; Nasir et al., 2023). In recent decades, neuroimaging data have grown, allowing ML and DL algorithms to better characterize AD. The authors used such methodologies to offer prospective AD diagnosis and prognostic outcomes (Nagarajan et al., 2021). These works executed features from several image processing pipeline streams using random forest classifier, decision tree, or support vector machine (SVM). Lately, DL techniques have showed potential in medical imaging with good picture classification accuracy (Ajagbe et al., 2021). Automatic feature extraction from images using CNNs and transfer learning (TL) is more efficient than typical ML methods (Raju et al., 2021). However, working with medical data is problematic due to imbalanced dataset, including AD. In this strategy, various sample sizes are used for different classes, the model is always biased, and it cannot generalize beyond the training dataset. DL models can process raw data better than simple feed forward, but they can overfit when solving complicated problems such as class imbalance. In real-world circumstances, such models perform poorly in generalization, efficacy, and reliability. The main contribution of this study is as follows:

- We introduce a contrast enhancement method called haze-reduced local-global, inspired by the haze reduction principle.
- We suggest a new global CNN-transformer architecture, InGSA, for the classification of multiclass AD. A pre-trained CNN is integrated with a specialized transformer module in InGSA network.
- This network, comprised of several generalized self-attention module (GSA), is designed to effectively capture extensive feature dependencies across different brain regions by establishing global connections along both the channel and spatial dimensions.
- The InGSA model is tested on a two publicly available dataset, where we also use various pre-trained CNN models to demonstrate its effectiveness. Furthermore, we perform a comparative analysis between InGSA and modern attention mechanisms, as well as the latest approaches in multiclass AD classification.

The structure of this research is comprehensively examined in the following manner: Related works are detailed in Section 2. Section 3 delineates the fundamental concepts and proposed methodology. The experimental results are the subject of Section 4. The study is concluded in Section 5.

2 Related work

Over the past few years, the usage of DL methods for the identification of AD has received much attention (Mohammed et al., 2021; Ahmed et al., 2022; Menagadevi et al., 2023). For instance, a study employed DL with stacked auto-encoders and uses the softmax function in the final layer to address problem of bottlenecks. Their approach needed far less training data compared to their peers, as well as very small input to classify several groups with ∼87.70% accuracy. One of the observations from the current study was that the use of several features improves classification (Frizzell et al., 2022). Furthermore, a classification framework was built, based on the use of multiple different input databases since it is complementary. To combine features from different modalities, they used a process known as non-linear graph mixture model. Using this method, the areas under the curve were calculated with 98.1% accuracy when differentiating between AD and CN images, 82.40% between NC and MCI images, with the overall classification performance being 77.90% (Guo and Zhang, 2020).

A novel rapid, low-cost, and efficient diagnostic model was implemented using brain MRI scans. They used DenseNet121 model which is a computationally heavy model, and to this model, they achieved an accuracy of 87% in detecting the disease. To rectify this, the authors employed an idea of fine-tuning two models of AlexNet and LeNet models where features were extracted in three ways through parallel filters. The new model they came up with was able to predict the disease with an accuracy of 93% (Hazarika et al., 2023). In the same manner, the researchers in Acharya et al. (2021) used VGG-16 based CNN transfer learning to diagnose AD with an overall accuracy of 95.7%. Another study used DL for distinguishing dementia and Alzheimer's from the MRI images (Murugan et al., 2021).

Abbreviations: CNN, convolutional neural network; HRLG, haze-reduced local-global; GSA, generalized self-attention; AD, Alzheimer's disease; MCI, moderate cognitive impairment; EMCI, Early MCI; LMCI, late MCI; MRI, magnetic resonance imaging; PET, positron emission tomography; ML, machine learning; DL, deep learning; SVM, support vector machine; TL, transfer learning; ELM, extreme learning machine; DAG, directed acyclic graph; Mob, MobileNet; Den, DenseNet201; Res, ResNet50; Sq, SqueezeNet; InV3, inceptionV3; CBAM, convolutional block attention module; CSDAB, channel split dual attention block; ViT, vision transformer; DEiT, data efficient image transformer; PVT, pyramid vision transformer.

The approach used in Murugan et al. (2021) learns individual Alzheimer's likelihood using multilayer perceptron representations and also generates disease probability heat maps from brain region activity. To overcome the problem of class imbalance, the samples are divided in equal proportion. The five ADNI subtypes consist of 1,296 images comprising of AD, MCI, EMCI, LMCI, and CN images processing the DEMNET model by resizing the images to 176 × 176 and obtained an accuracy of 84.83%. In the same way, Oktavian et al. (2022) presented the fine-tuned ResNet18 model for distinguishing between MCI, AD, and CN using MRI and PET datasets. This model incorporated transfer learning and used the technologies such as weighted loss function for ascending the class imbalance, and mish activation function to augment its accuracy, and it obtained 88.3% overall classification. On the other hand, the authors in Dyrba et al. (2021) adopted a CNN with 663 T1-weighted MRI scans belonging to dementia and amnestic MCI patients. To confirm their model, they performed cross-validation and used an additional three datasets that included an overall of 1,655 cases. To further provide the clinical relevance of the method, they correlated the relevance scores to the hippocampal volume. A friendly model assessment tool was created through importance maps of 3D CNN, achieving accuracy of 94.9% of AD vs. CN. A particular drawback of many papers on the detection of Alzheimer's is related to the imbalance of classes, which creates problems of overfitting and lowering predictive ability in almost all existing deep learning models. The yield is further magnified by the fact that realistic training data are also scarce. To overcome this, we utilized the data augmentation approach to balance datasets and improve DL results since the technique synthesizes new data samples.

3 Proposed methodology

The configuration of the proposed InGSA is illustrated in Figure 1, comprising a fine-tuned CNN model, a generalized self-attention (GSA), and a classifier. The fine-tuned CNN models aid in extracting abstract feature representations from the input MR images. The GSA block has various components to comprehend global interdependence across spatial and channel dimensions, facilitating the extraction of more nuanced and category-specific information. The extreme learning machine (ELM) classifier is employed to categorize AD. This section offers a comprehensive overview of the InGSA architecture and its fundamental components.

3.1 Haze reduced local global image enhancement

Traditional haze elimination procedures are developed for improving the visual distinctiveness of scenes by increasing the contrast and color saturation. By applying these techniques, the total clarity of the scene which is captured in the given image is likely to be enhanced. In this research, we formally propose a new type of contrast enhancement method that adopts both haze removal and local-global transformation techniques.

Let D denotes a complex image database that is composed of N images. While the original image is represented by the dimensions of $N \times M \times 3$ as $I(x, y)$, the $Y(x, y)$ denotes the improved image. First, a haze reduction method utilizing the dark channel prior is employed on the first image. This process of haze reduction can be mathematically expressed as follows:

$$C(x) = \gamma(x)j(x) + l(1 - t(x)) \tag{1}$$

where C denotes the measured intensity values, γ represents the scene radiance, $j(x)$ designates the transmission map, and l denotes the atmospheric light intensity. The dehazing algorithm utilized aims to restore the scene radiance γ based on the estimations of both the transmission map and the atmospheric light, as expressed in the following manner:

$$\gamma(x) = \frac{C(x) - \alpha}{\max(t(x), t_0)} + \alpha \tag{2}$$

The resulting $\gamma(x)$ is subsequently employed to calculate the global contrast of an image using the following equation:

$$G_0 = (1 + g_k) \times (G_i - k_{\text{mean}}) + \sigma \tag{3}$$

In this regard, G_0 stands for the global contrast image of the original image while g_k represents gain factor of global contrast, G_i for the value of pixel $\gamma(x)$, k_{mean} for the overall average pixel value of $\gamma(x)$, and σ for the standard deviation of $\gamma(x)$. In the subsequent step, we assessed the local contrast of the haze-reduced image using the following mathematical expression:

$$H(x, y) = \frac{LC}{\sigma(i, j) + \varphi} \times \mu(x, y) \tag{4}$$

where LC for local contrast, φ for a small constant, and $\mu(x, y)$ means the mean value of the dehazed image. Finally, these two resultant images of local and global contrast were incorporated toward a single image in this way that we adopt the following mathematical formula to produce the final enhancement output.

$$Y(x, y) = [G(x, y) + H(x, y)] - I(x, y)] \tag{5}$$

3.2 Deep transfer learning

InceptionV3 is a directed acyclic graph (DAG) network that has 316 layers and 350 links that include 94 as convolutional layers (Szegedy et al., 2016). Such a structure facilitates the provision of adequate employment of complicated dependency relations in the network, having many inputs and outputs at different layers. Differently from the standard CNN model in which the filter size is fixed throughout the layers of that model, InceptionV3 has different filter sizes within the same layer, which increases its capability of feature extraction on the data. Originally trained on ImageNet (Deng et al., 2009) on which includes over one million images split into one thousand categories, the InceptionV3 has the ability to read most features. The model takes images of input size 299 × 299 × 3. In this study, the model has been adapted to classify various stages of AD for which transfer learning from the ImageNet

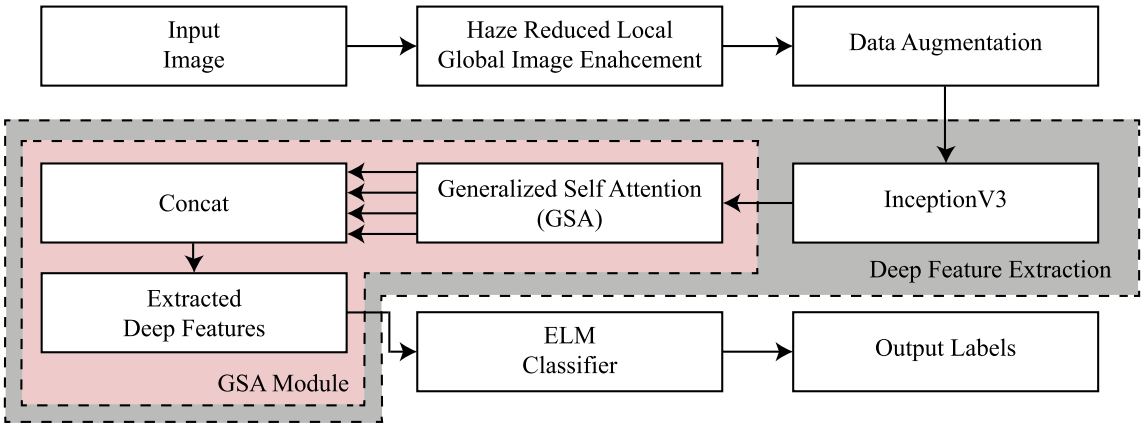


FIGURE 1
General flow of the proposed framework.

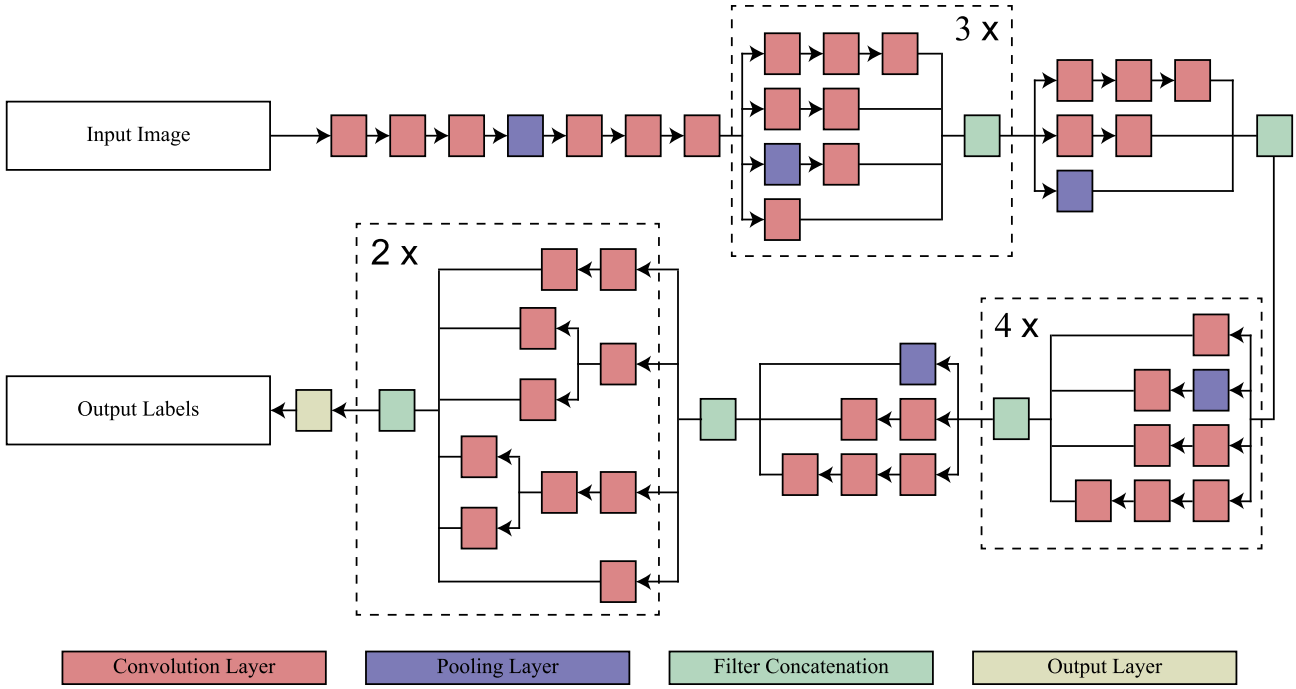


FIGURE 2
Architecture of InceptionV3 model.

training phase has been applied to achieve efficient medical image classification. The architecture of InceptionV3 model is shown in Figure 2.

Transfer learning (Pan and Yang, 2009) is a popular method in recognition and detection tasks, allowing for improved model performance by leveraging pre-trained models. In this context, the domain D consists of a feature vector $Y = y_1, y_2, \cdots, y_n$ with a corresponding probabilistic distribution $P(Y)$, forming $B = Y, P(Y)$. The task, denoted as T, consists of the ground truth $Z = z_1, z_2, \cdots, z_n$. The function can be expressed in probabilistic form as $P(z|y)$. In the context of transfer learning, this can be represented concerning the source domain as $B_T = (x(T_1), x(T_2)), (x(T_2), x(T_2)), \cdots, (x(T_n), x(T_n))$ along with the learning rate S_T. The target output is denoted as $B_S = (x(S_1), x(S_2)), (x(S_2), x(S_2)) \cdots, (x(S_n), x(S_n))$, and the associated function for the targeted neural network is represented as S_S. The primary objective of transfer learning is to improve the learning rate for predicting the target object by utilizing the recognition function $F_S(.)$, which is informed by training on both B_T and B_S, where $B_T \neq B_S$ and $S_T \neq S_S$. Inductive transfer learning proves to be effective in pattern recognition tasks. An annotated dataset is essential for efficient training and evaluation when implementing inductive transfer learning. This process can involve distinct class labels $Z_T \neq Z_S$ and differing distributions $P(Z_T|Y_T) \neq (Z_S|Y_S)$.

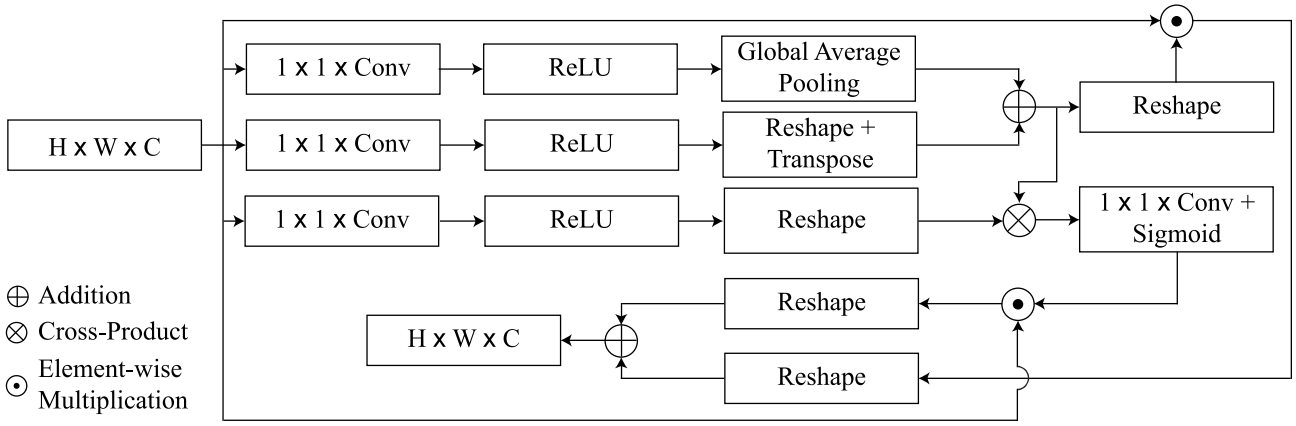


FIGURE 3
Overview of the GSA block.

3.3 Generalized self-attention module

The proposed GSA module is aimed at achieving detailed description of AD characteristics while avoiding irrelevant features. Its architecture was influenced by the self-attention mechanisms employed in GCNet (Cao et al., 2019; Zhang et al., 2019), as illustrated in Figure 3. However, unlike these methods, it positions global dependency across both spatial and channel dimensions at the same time. Spatial attention worked for the relationships of the global features in the spatial location, while channel attention worked on the importance of a point channel out of all the channels.

As initial input for the GSA module, we utilize the high level activation maps $Z \in \mathbb{R}^{H\times W\times C}$, whereas the GSA module returns the refined feature maps $Z_{gs} \in \mathbb{R}^{H\times W\times C}$. The feature map is divided into the keys, queries, and values, similar to a transformer architecture, which is supported by three attributes Q, K, and V.

The query function $q(Z)$ is defined by a convolution of 1×1 consisting of $C' = C/8$ channels and global average pooling to attain the vector $Q(Z) \in \mathbb{R}^{1\times C'}$. On the other hand, the key and value functions are carried out by 1×1 convolution followed by reshape operations but without global average pooling and the outputs are maps $K(Z) \in \mathbb{R}^{HW\times C'}$ and $V(Z) \in \mathbb{R}^{HW\times C'}$). Next, The spatial attention weights are generated by calculating the matrix product between Q and K and applying a softmax activation function given as

$$Z' = \phi(q(Z) \otimes k(Z)^T) \tag{6}$$

With regard to the abbreviations used here, we have $\otimes$ indicating the cross-product of the matrix, ϕ which stands for the softmax activation function of the formula and the double dagger T showing the operation of matrix transposition. Following this, the spatial attention feature map $Z_{sp} \in \mathbb{R}^{H\times W\times C}$ derived by performing element-wise multiplication among Z' and Z is as follows:

$$Z_{sp} = \text{reshape}(Z') \odot Z \tag{7}$$

Similarly, a matrix cross-product of Z' with $v(Z)$ leads to the channel attention weights which are passed through a 1×1 convolution layer and a sigmoid non-linearity. It also increases the channels from C' to C. This process is also called linear embedding. The mathematical formulation for this global transformation is given by

$$Z'' = \sigma(\text{conv}(Z' \otimes v(Z))) \tag{8}$$

Next, the channel-wise attention maps $Z_{ch} \in \mathbb{R}^{H\times W\times C}$ are calculated as

$$Z_{ch} = Z'' \odot Z \tag{9}$$

Finally, we integrate the spatial attention feature maps Z_{sp} and the channel attention maps Z_{ch} by taking their weighted sum, producing the refined attention feature map $Z_{gs} \in \mathbb{R}^{H\times W\times C}$, defined as

$$Z_{gs} = W_1 Z_{sp} + W_2 Z_{ch} \tag{10}$$

where W_1 and W_2 are two trainable scalar weights. In summary, GSA obtains the channel-wise and spatial dependencies concurrently from MR images and then improves the features representation. We merge the feature attention maps produced by the GSA heads through concatenation, preceding 1×1 convolution to generate the final output of the proposed GSA, denoted as $Z_{tm} \in \mathbb{R}^{H\times W\times C}$. Mathematically, $Z_t m$ can be represented as follows:

$$Z_{tm} = \text{concat}(Z^1_{gs}, Z^2_{gs}, \ldots, Z^h_{gs}) \tag{11}$$

In this study, h, representing the number of GSA heads, is set empirically to a value of 4. This specific choice of $h = 4$ was determined empirically.

3.4 Classification

The ELM [42] was used as a classifier to differentiate AD stages. Given z sample (Z, o), the ELM's output with no errors can be mathematically expressed as follows:

$$o = \sqrt{\sum \alpha t(w_i + b)} \tag{12}$$

In this instance, the activation function is denoted by $t(\cdot)$, and the input and output samples are Z and o, respectively. The variables w and b are weights and bias, respectively, and α is the weight coefficient. The output O is provided as $O = H\alpha$ whereby $O = (o_1, o_2, \cdots, o_n)$ symbolizes the output vector and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ denotes the weight vector. The hidden layers can be expressed as

$$H = \begin{bmatrix} t(w_1 i_1 + b_1) & \cdots & t(w_n i_1 + b_n) \\ \vdots & \ddots & \vdots \\ t(w_1 i_m + b_1) & \cdots & t(w_n i_m + b_n) \end{bmatrix} \tag{13}$$

The number of nodes in the hidden layer needs to be below the total amount of samples. Description of the structured model of a single hidden-layer ELM neural network utilized for AD classification is provided in Equation 13. The hidden layer, denoted by H, is composed of nodes and activation functions $t(\cdot)$. Weights w_i and biases b are connected to each hidden layer node, with i ranging from 1 to m and representing input variables. The formula used in the production of the output of the hidden layer O is the summation of the product between each of the node's activation function and weights then passed through $t(\cdot)$. The mechanism can be mathematically represented as $O = H\alpha$. Equation 13 defines the structure of the hidden layer, which further elucidates that the H is a concatenation of n nodes. The weighted sum of input features $i = (i_1, i_2, \cdots, i_m)$ is computed for each node's activation by utilizing weights w_i and biases b for each node. The function $t(\cdot)$ adds non-linearity and provides the network with ability to learn more complex input data patterns. Determining the quantity of nodes in the hidden layers is essential; they should be fewer than the amount of samples to avert overfitting. During training, we obtain the weights and biases to minimize the mapping function between the input features and the associated output for AD classification.

4 Experimental results

The analysis and experimental results of the proposed models are detailed in this section. The presentation includes information regarding the dataset, implementation characteristics, and comparison analysis.

4.1 Experimental setup and dataset

The model was trained on a high-performance machine equipped with an Intel Core i9-14900HX processor and an NVIDIA RTX 4090 GPU, providing substantial computational power for deep learning tasks. The system included 64GB of DDR5 RAM operating at 5,600MT/s, ensuring efficient handling of large-scale data. CUDA 12.6 was utilized to enable GPU-accelerated training. The model was trained with a learning rate of 0.0001, a value chosen to balance the stability and convergence speed of the training process.

TABLE 1 ADNI dataset image count before and after augmentation.

Classes	No. of images	Augmented	No. of utilized images
AD	8,346	n/a	500
CN	8,650	n/a	500
EMCI	480	n/a	480
LMCI	144	432	432
MCI	1,155	n/a	500

TABLE 2 Total number of images and number of images utilized from OASIS dataset.

Classes	No. of images	No. of utilized images
Mild dementia	5,002	500
Moderate dementia	488	488
Non-demented	67,200	500
Very mild dementia	13,700	500

In this experiment, ADNI dataset was used which consisted of five classes: AD, CN, EMCI, LMCI, and MCI. The original number of images varied significantly across classes, with AD, CN, and MCI having thousands of images, while EMCI and LMCI had considerably fewer as shown in Table 1. To address this class imbalance, data augmentation was applied exclusively to the LMCI class, which originally had only 144 images. Through augmentation, the LMCI class was expanded to 432 images, increasing the total number of samples used in the training process. For the other classes, 500 images were randomly selected from AD and CN, while all available images were used for EMCI and MCI.

The augmentation methods used to enhance the LMCI class included rotation, scaling, and flipping. Rotation involved rotating images by various angles to introduce diversity without altering key signal features. Scaling was applied to adjust the size of images while maintaining their aspect ratio, simulating variability in data capture. Flipping, both horizontally and vertically, was also used to further diversify the dataset, making the model more robust to orientation changes. These augmentation techniques were critical in improving class balance and ensuring better generalization during model training.

Another dataset used in this experiment is the Open Access Series of Imaging Studies (OASIS), a widely utilized resource for neuroimaging research, particularly in the study of brain health and dementia. Table 2 presents the distribution of images from the OASIS dataset across four classes: Mild Dementia, Moderate Dementia, Non-Demented, and Very Mild Dementia.

4.2 ADNI results

Table 3 indicates the performances of the proposed model in terms of classification for Alzheimer's detection under the different cognitive conditions. The network can accurately predict an image with an average of 96.67% and high values of precision, recall, and F1-scores for all classes. For the LMCI class, the study realized a precision of 98.80%, while at the same time, AD has the highest recall of 98.00% to show effective detection. All classes have an F1-score of more than 96%, and it can be seen that the approach balanced precision and recall. The AUC values are also high, and AD reached 98.21%. Confusion matrix for proposed model is given in Figure 4.

Table 4 gives a quantitative analysis of the number of correct detections of the various models augmented with different

TABLE 3 Classification performance of InGSA on ADNI dataset.

Class	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
AD	94.23	98.00	96.08	98.21
CN	97.92	95.92	96.91	97.70
EMCI	95.88	96.88	96.37	97.92
LMCI	98.80	95.35	97.04	97.55
MCI	97.00	97.00	97.00	98.11
Accuracy	96.67			
Macro average	97.76	96.63	96.68	97.90
Weighted average	96.71	96.67	96.67	97.91

TABLE 4 Performance comparison of different attention mechanisms.

Attention mechanism	Accuracy (%)				
	Mob	Den	Res	Sq	InV3
None	82.76	76.65	71.27	86.32	90.17
Self attention	85.43	79.89	76.39	87.36	93.44
CBAM	83.70	78.72	71.89	89.78	91.23
CSDAB	83.00	78.21	73.47	88.24	90.87
Proposed (GSA)	84.54	82.87	79.98	91.43	96.67

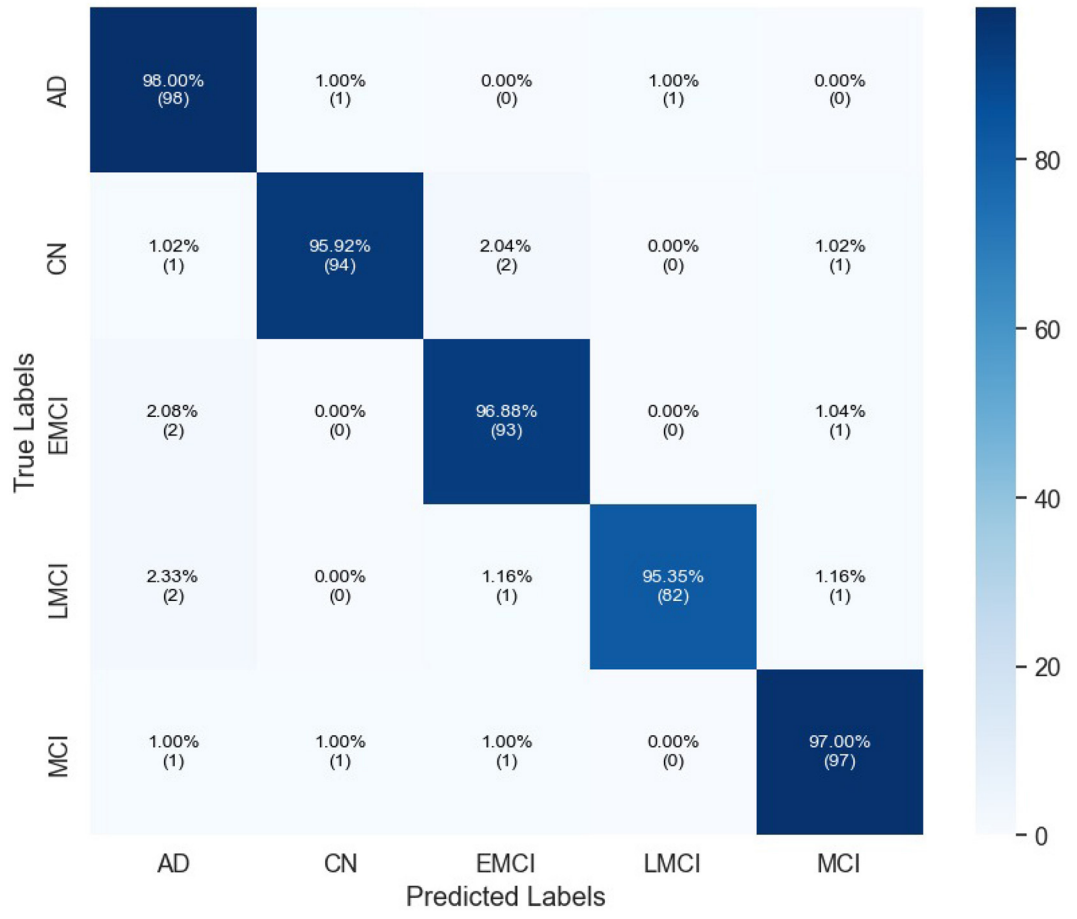


FIGURE 4
Confusion matrix of proposed InGSA on ADNI Dataset.

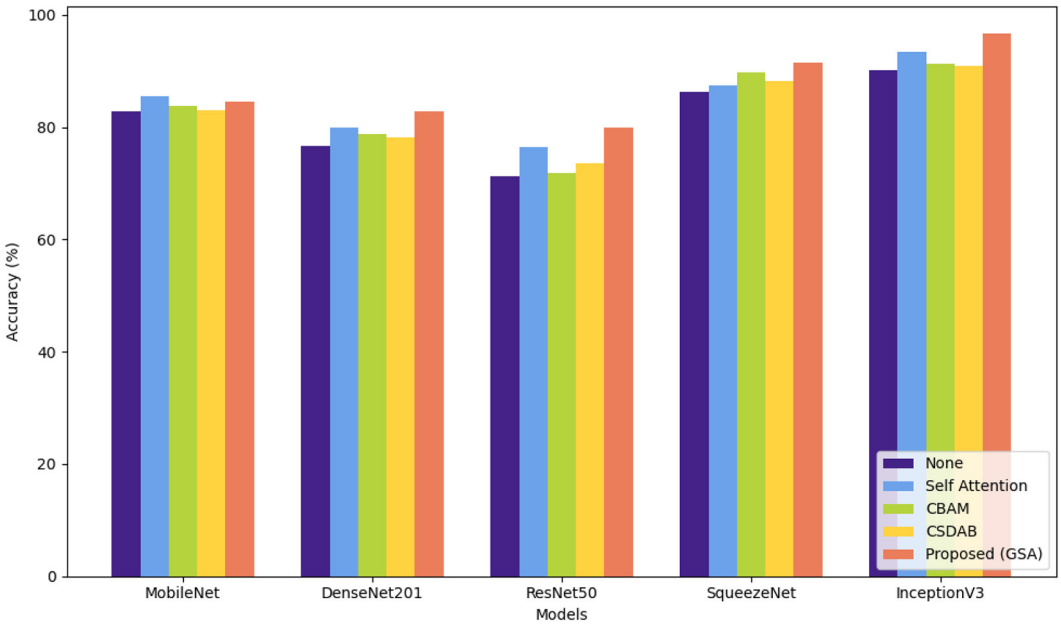


FIGURE 5
Visual comparison of various attention mechanisms on ADNI dataset.

attention mechanisms: MobileNet (Mob), DenseNet201 (Den), ResNet50 (Res), SqueezeNet (Sq), and InceptionV3 (InV3). With the proposed GSA, the largest performance improvements were achieved with InceptionV3, from 90.17% to 96.67% (with attention) and with SqueezeNet from 86.32% to 91.43%. Here, DenseNet201 shows improvement of 6.22% from 76.65 to 82.87, while ResNet50 goes from 71.27% to 79.98%. Self-Attention (Zhang et al., 2019) also presented substantial enhancements for InceptionV3 from 93.15% to 93.44%, as well as for SqueezeNet from 86.59% to 87.36%. Both convolutional block attention module (CBAM) (Woo et al., 2018) and channel split dual attention block (CSDAB) (Dutta and Nayak, 2022) result in moderate accuracy increases, with CBAM improving SqueezeNet to 89.78% and CSDAB raising it to 88.24%. In general, GSA consistently improves accuracy in all models being tested. Visual analysis of attention mechanisms with pre-trained model on ADNI dataset is shown in Figure 5.

Table 5 shows accuracy and F1-score when comparing the current existing models, namely, vision transformer (ViT), data efficient image transformer (DEiT), and pyramid vision transformer (PVT) with the InGSA model proposed in this study. Out of all the models, the DEiT comes with the highest accuracy and F1-score with accuracies of 89.44%, and F1-score of 88.12%, with PVT coming second with accuracies of 88.48% and F1-scores of 85.98%. ViT has the worst performance with an accuracy of 86.24% and the F1-score at 84.67%. Comparing the proposed model InGSA with others, it is clear that the proposed model InGSA outperforms the other models with accuracy of 96.67% and F1-score of 96.68% that indicate effectiveness of the proposed model InGSA.

TABLE 5 Comparison of proposed InGSA with transformer-based models on ADNI dataset.

Model	Accuracy (%)	F1-score (%)
ViT	86.24	84.67
DEiT	89.44	88.12
PVT	88.48	85.98
InGSA	96.67	96.68

TABLE 6 Classification performance of InGSA on OASIS dataset.

Class	Precision (%)	Recall (%)	F1-score (%)	AUC (%)
Mild demented	98.02	99.00	98.51	99.16
Moderate demented	98.97	97.96	98.46	98.81
Non-demented	97.94	95.00	96.45	97.16
Very mild demented	97.09	100.00	98.52	99.50
Accuracy	97.99			
Macro average	98.00	97.99	97.98	98.66
Weighted average	98.00	97.99	97.98	98.66

4.3 OASIS results

Table 6 presents the classification performance for different categories of dementia using OASIS dataset. High precision of

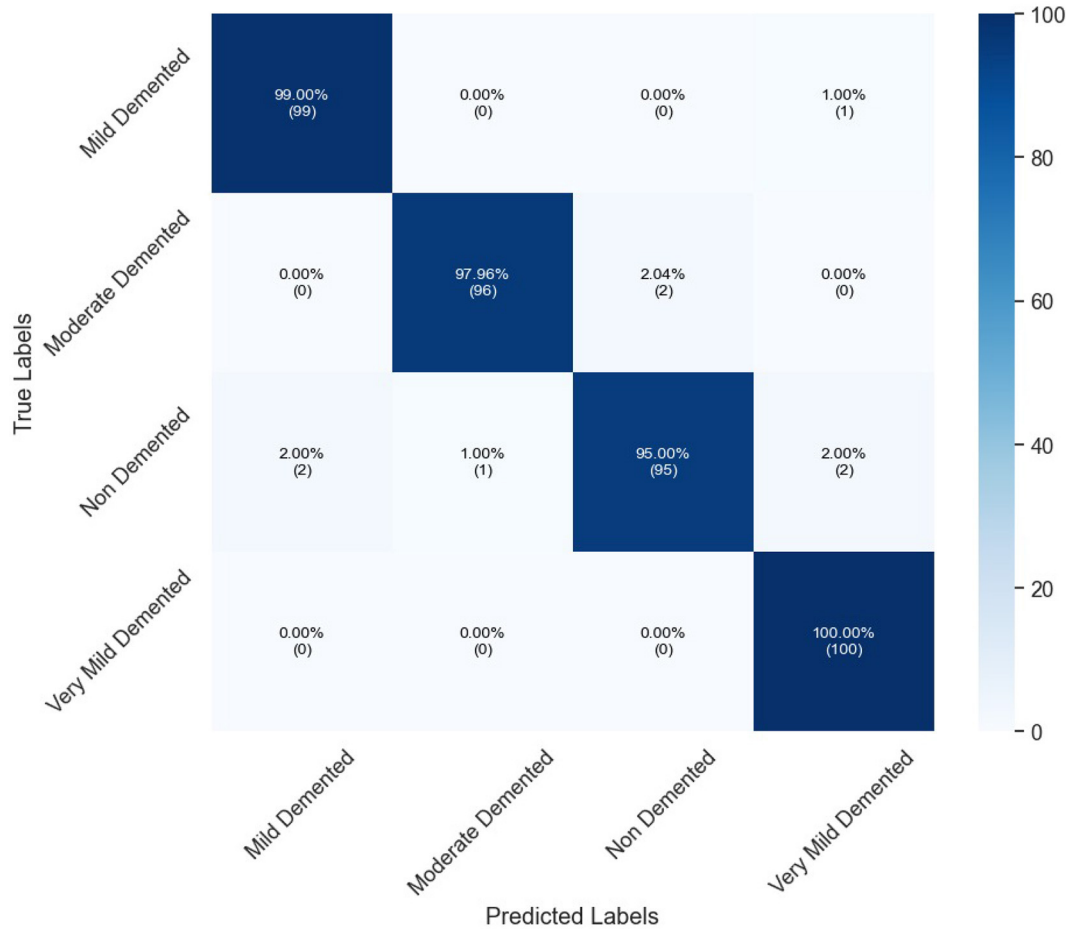


FIGURE 6
Confusion matrix of proposed InGSA on OASIS dataset.

TABLE 7 Performance comparison of different attention mechanisms.

Attention mechanism	Accuracy (%)				
	Mob	Den	Res	Sq	InV3
None	87.35	80.78	72.34	87.54	92.87
Self attention	88.45	84.67	73.67	89.79	94.76
CBAM	88.89	84.32	77.93	88.38	93.62
CSDAB	88.95	82.19	74.05	88.75	93.90
Proposed (GSA)	90.74	87.64	79.85	91.02	97.99

all classes set by the model, specifically Moderate Demented presenting 98.97% and Mild Demented 98.02%. The Recall is exceptional for Very Mild Demented at 100% which means that all the cases belonging to this class are identified rightly. The F1-score, therefore, averaged over all classes is unbiased, being 96.45% for Non-Demented and 98.52% for both Mild and Very Mild Demented classes. AUC's are high, notably Very Mild Demented with a highest of 99.50%. In general, the proposed model renders high performance in the presented study with the overall accuracy of 97.99%. Figure 6 depicts confusion matrix of proposed model for OASIS dataset.

Table 7 shows the percentage of classification accuracy of multiple models when using the OASIS dataset, with different attention mechanisms. All the attention approaches improve on the baseline accuracy of all the models including the proposed GSA model. For instance, they achieve 90.74% accuracy with MobileNet and an outstanding 97.99% with InceptionV3 confirming how efficient the proposed approach is in enhancing model accuracy. Self-attention mechanism also plays a useful role, especially in MobileNet and InceptionV3 models and in this experiment reached a throughput of 88.45 and 94.76%, correspondingly. While structure imported with CBAM and CSDAB mechanisms may be less fortunate than the GSA model, it has revealed improvements. Figure 7 illustrates the visual analysis of attention mechanisms using a pre-trained model on the OASIS dataset.

The findings in Table 8 reveal that all algorithms provide high accuracy, where proposed InGSA model performs the best with accuracy of 97.99% and F1-score of 97.88%. This implies that InGSA is proud not only to classify instances accurately but also to have a low percentage of false positive and false

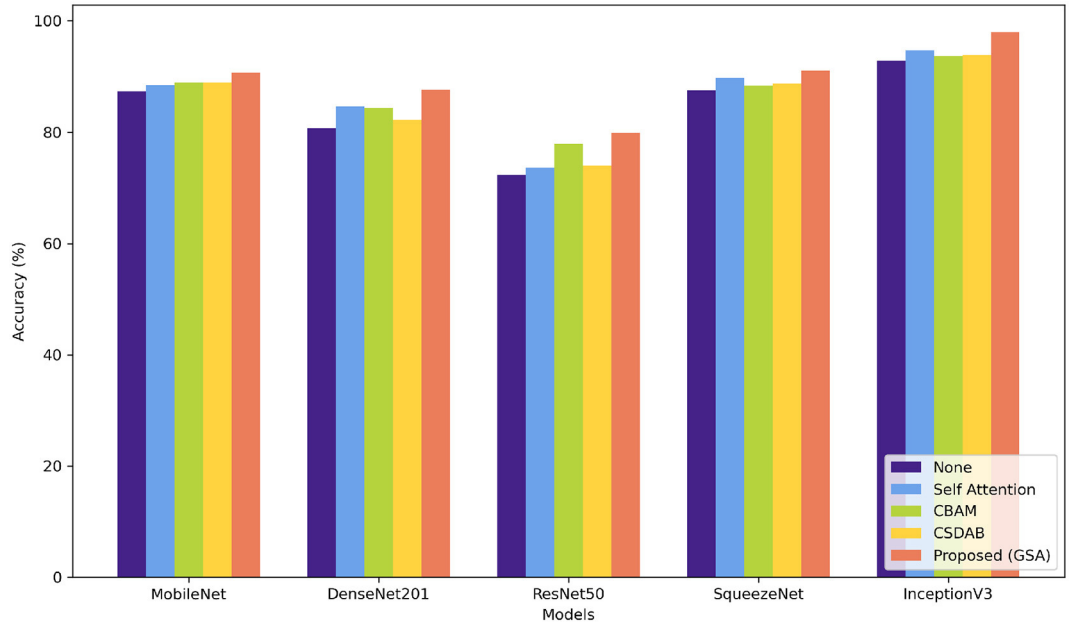


FIGURE 7
Visual comparison of various attention mechanisms on OASIS dataset.

TABLE 8 Comparison of proposed InGSA with transformer-based models on OASIS dataset.

Model	Accuracy (%)	F1-score (%)
ViT	91.65	86.78
DEiT	89.43	89.12
PVT	93.79	92.61
InGSA	97.99	97.88

negative. Next is the PVT model which gives classification accuracy of 93.79% and F1-score of 92.61% demonstrating the good classification prowess of the model. The performance of both ViT and DEiT models is reasonable, with accuracies of 91.65% and 89.43%, respectively.

5 Conclusion

Alzheimer's disease, diagnosed and classified with multiclass datasets in the early stage, needed a proficient automatic system identification. This study puts forward a CNN-Transformer model to diagnose Alzheimer's cases from multiclass datasets using transfer learning. First, a method of contrast enhancement is utilized to help better visualize important features. Furthermore, we introduce a new global CNN-transformer network known as InGSA for multiclass AD classification to facilitate end-to-end training. The InGSA architecture is based on the CNN and transformer, and GSA blocks are placed on top of pre-trained InceptionV3 model. GSA blocks are important for expression subscalar detection of global dependencies of features. The GSA component improves the extraction of detailed information by learning channel-wise and spatial-wise attention weights at the same time. In-depth experiments on two benchmark datasets demonstrate that our proposed InGSA achieves superior performance compared to the state-of-the-art techniques. Furthermore, GSA yields better results than other traditional attention methods. For the future works, we aim to test GSA on more extensive set and diverse dataset, and we also want to apply our proposed method in the other vision-related tasks.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://adni.loni.usc.edu/data-samples/adni-data/.

Author contributions

FB: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Writing – original draft. AA: Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors would like to express sincere gratitude to the Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Saudi Arabia, for their invaluable support and guidance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acharya, H., Mehta, R., and Singh, D. K. (2021). "Alzheimer disease classification using transfer learning," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)* (Erode: IEEE), 1503–1508. doi: 10.1109/ICCMC51019.2021.9418294

Ahmed, G., Er, M. J., Fareed, M. M. S., Zikria, S., Mahmood, S., He, J., et al. (2022). Dad-net: classification of Alzheimer's disease using ADASYN oversampling technique and optimized neural network. *Molecules* 27:7085. doi: 10.3390/molecules27207085

Ajagbe, S. A., Amuda, K. A., Oladipupo, M. A., Oluwaseyi, F. A., and Okesola, K. I. (2021). Multi-classification of alzheimer disease on magnetic resonance images (MRI) using deep convolutional neural network (DCNN) approaches. *Int. J. Adv. Comput. Res.* 11:51. doi: 10.19101/IJACR.2021.1152001

Böhle, M., Eitel, F., Weygandt, M., and Ritter, K. (2019). Layer-wise relevance propagation for explaining deep neural network decisions in MRI-based Alzheimer's disease classification. *Front. Aging Neurosci.* 11:456892. doi: 10.3389/fnagi.2019.00194

Cao, Y., Xu, J., Lin, S., Wei, F., and Hu, H. (2019). "GCNET: non-local networks meet squeeze-excitation networks and beyond," in *Proceedings of the IEEE/CVF international conference on computer vision workshops* (Seoul: IEEE). doi: 10.1109/ICCVW.2019.00246

Choi, B.-K., Madusanka, N., Choi, H.-K., So, J.-H., Kim, C.-H., Park, H.-G., et al. (2020). Convolutional neural network-based MR image analysis for Alzheimer's disease classification. *Curr. Med. Imaging* 16, 27–35. doi: 10.2174/1573405615666191021123854

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L., et al. (2009). "Imagenet: a large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition* (Miami, FL: IEEE), 248–255. doi: 10.1109/CVPR.2009.5206848

Dutta, T. K., and Nayak, D. R. (2022). "CDANET: channel split dual attention based CNN for brain tumor classification in MR images," in *2022 IEEE international conference on image processing (ICIP)* (Bordeaux: IEEE), 4208–4212. doi: 10.1109/ICIP46576.2022.9897799

Dyrba, M., Hanzig, M., Altenstein, S., Bader, S., Ballarini, T., Brosseron, F., et al. (2021). Improving 3d convolutional neural network comprehensibility via interactive visualization of relevance maps: evaluation in Alzheimer's disease. *Alzheimers Res. Ther.* 13, 1–18. doi: 10.1186/s13195-021-00924-2

Edmonds, E. C., McDonald, C. R., Marshall, A., Thomas, K. R., Eppig, J., Weigand, A. J., et al. (2019). Early versus late MCI: improved mci staging using a neuropsychological approach. *Alzheimers Dement.* 15, 699–708. doi: 10.1016/j.jalz.2018.12.009

Frizzell, T. O., Glashutter, M., Liu, C. C., Zeng, A., Pan, D., Hajra, S. G., et al. (2022). Artificial intelligence in brain MRI analysis of Alzheimer's disease over the past 12 years: a systematic review. *Ageing Res. Rev.* 77:101614. doi: 10.1016/j.arr.2022.101614

Gao, F. (2021). Integrated positron emission tomography/magnetic resonance imaging in clinical diagnosis of Alzheimer's disease. *Eur. J. Radiol.* 145:110017. doi: 10.1016/j.ejrad.2021.110017

Gunawardena, K., Rajapakse, R., and Kodikara, N. D. (2017). "Applying convolutional neural networks for pre-detection of Alzheimer's disease from structural MRI data," in *2017 24th international conference on mechatronics and machine vision in practice (M2VIP)* (Auckland: IEEE), 1–7. doi: 10.1109/M2VIP.2017.8211486

Guo, H., and Zhang, Y. (2020). Resting state fMRI and improved deep learning algorithm for earlier detection of Alzheimer's disease. *IEEE Access* 8, 115383–115392. doi: 10.1109/ACCESS.2020.3003424

Hazarika, R. A., Maji, A. K., Kandar, D., Jasinska, E., Krejci, P., Leonowicz, Z., et al. (2023). An approach for classification of Alzheimer's disease using deep neural network and brain magnetic resonance imaging (MRI). *Electronics* 12:676. doi: 10.3390/electronics12030676

Kang, L., Jiang, J., Huang, J., and Zhang, T. (2020). Identifying early mild cognitive impairment by multi-modality MRI-based deep learning. *Front. Aging Neurosci.* 12:206. doi: 10.3389/fnagi.2020.00206

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., Stadlan, E. M., et al. (1984). Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group* under the auspices of department of health and human services task force on Alzheimer's disease. *Neurology* 34, 939–939. doi: 10.1212/WNL.34.7.939

Menagadevi, M., Mangai, S., Madian, N., and Thiyagarajan, D. (2023). Automated prediction system for Alzheimer detection based on deep residual autoencoder and support vector machine. *Optik* 272:170212. doi: 10.1016/j.ijleo.2022.170212

Mohammed, B. A., Senan, E. M., Rassem, T. H., Makbol, N. M., Alanazi, A. A., Al-Mekhlafi, Z. G., et al. (2021). Multi-method analysis of medical records and MRI images for early diagnosis of dementia and Alzheimer's disease based on deep learning and hybrid methods. *Electronics* 10:2860. doi: 10.3390/electronics10222860

Murugan, S., Venkatesan, C., Sumithra, M., Gao, X.-Z., Elakkiya, B., Akila, M., et al. (2021). Demnet: a deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. *IEEE access* 9, 90319–90329. doi: 10.1109/ACCESS.2021.3090474

Nagarajan, S. M., Muthukumaran, V., Murugesan, R., Joseph, R. B., and Munirathanam, M. (2021). Feature selection model for healthcare analysis and classification using classifier ensemble technique. *Int. J. Syst. Assur. Eng. Manag.* 1–12. doi: 10.1007/s13198-021-01126-7

Nasir, I. M., Bibi, A., Shah, J. H., Khan, M. A., Sharif, M., Iqbal, K., et al. (2021). Deep learning-based classification of fruit diseases: an application for precision agriculture. *Comput. Mater. Contin.* 66, 1949–1962. doi: 10.32604/cmc.2020.012945

Nasir, I. M., Khan, M. A., Yasmin, M., Shah, J. H., Gabryel, M., Scherer, R., et al. (2020). Pearson correlation-based feature selection for document classification using balanced training. *Sensors* 20:6793. doi: 10.3390/s20236793

Nasir, I. M., Raza, M., Shah, J. H., Khan, M. A., Nam, Y.-C., Nam, Y., et al. (2023). Improved shark smell optimization algorithm for human action recognition. *Comput. Mater. Contin.* 76, 2667–2684. doi: 10.32604/cmc.2023.035214

Nasir, I. M., Raza, M., Shah, J. H., Wang, S.-H., Tariq, U., Khan, M. A., et al. (2022). Harednet: A deep learning based architecture for autonomous video surveillance by recognizing human actions. *Comput. Electr. Eng.* 99:107805. doi: 10.1016/j.compeleceng.2022.107805

Nozadi, S. H., Kadoury, S., and The Alzheimer's Disease Neuroimaging Initiative (2018). Classification of Alzheimer's and mci patients from semantically parcelled pet images: a comparison between av45 and FDG-pet. *Int. J. Biomed. Imaging* 2018:1247430. doi: 10.1155/2018/1247430

Oktavian, M. W., Yudistira, N., and Ridok, A. (2022). Classification of Alzheimer's disease using the convolutional neural network (CNN) with transfer learning and weighted loss. *arXiv* [Preprint]. arXiv:2207.01584. doi: 10.48850/arXiv.2207.01584

Pan, S. J., and Yang, Q. (2009). A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191

Raju, M., Gopi, V. P., and Anitha, V. (2021). "Multi-class classification of Alzheimer's disease using 3dcnn features and multilayer perceptron," in *2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)* (Chennai: IEEE), 368–373. doi: 10.1109/WiSPNET51692.2021.9419393

Ramzan, F., Khan, M. U. G., Rehmat, A., Iqbal, S., Saba, T., Rehman, A., et al. (2020). A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. *J. Med. Syst.* 44, 1–16. doi: 10.1007/s10916-019-1475-2

Simic, G., Stanic, G., Mladinov, M., Jovanov-Milosevic, N., Kostovic, I., Hof, P. R., et al. (2009). Does Alzheimer's disease begin in the brainstem? *Neuropathol. Appl. Neurobiol.* 35, 532–554. doi: 10.1111/j.1365-2990.2009.01038.x

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (Las Vegas, NV: IEEE), 2818–2826. doi: 10.1109/CVPR.2016.308

Tehsin, S., Nasir, I. M., Damaševičius, R., and Maskeliūnas, R. (2024). DASAM: Disease and spatial attention module-based explainable model for brain tumor detection. *Big Data Cogn. Comput.* 8:97. doi: 10.3390/bdcc8090097

Tuvshinjargal, B., and Hwang, H. (2022). VGG-C transform model with batch normalization to predict Alzheimer's disease through MRI dataset. *Electronics* 11:2601. doi: 10.3390/electronics11162601

Varatharajah, Y., Ramanan, V. K., Iyer, R., and Vemuri, P. (2019). Predicting short-term mci-to-ad progression using imaging, CSF, genetic factors, cognitive resilience, and demographics. *Sci. Rep.* 9:2235. doi: 10.1038/s41598-019-38793-3

Woo, S., Park, J., Lee, J.-Y., and Kweon, I. S. (2018). "CBAM: convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*, eds. V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss (Cham: Springer), 3–19. doi: 10.1007/978-3-030-01234-2_1

Yousafzai, S. N., Shahbaz, H., Ali, A., Qamar, A., Nasir, I. M., Tehsin, S., et al. (2024). X-news dataset for online news categorization. *Int. J. Intell. Comput. Cybern.* 17, 737–758. doi: 10.1108/IJICC-04-2024-0184

Zhang, H., Goodfellow, I., Metaxas, D., and Odena, A. (2019). "Self-attention generative adversarial networks," in *International conference on machine learning* (Long Beach, CA: PMLR), 7354–7363.

Zhang, T., Zhao, Z., Zhang, C., Zhang, J., Jin, Z., Li, L., et al. (2019). Classification of early and late mild cognitive impairment using functional brain network of resting-state fMRI. *Front. Psychiatry* 10:572. doi: 10.3389/fpsyt.2019.00572

frontiers | Frontiers in Artificial Intelligence

TYPE Original Research
PUBLISHED 26 March 2025
DOI 10.3389/frai.2025.1426455

OPEN ACCESS

EDITED BY
Tony J. Prescott,
The University of Sheffield, United Kingdom

REVIEWED BY
Liliana Ibeth Barbosa Santillan,
University of Guadalajara, Mexico
Aung Htet,
Sheffield Hallam University, United Kingdom

*CORRESPONDENCE
Mohamad Eid
✉ mohamad.eid@nyu.edu

RECEIVED 24 May 2024
ACCEPTED 10 March 2025
PUBLISHED 26 March 2025

CITATION
Babushkin V, Alsuradi H, Al-Khalil MO and Eid M (2025) Analyzing handwriting legibility through hand kinematics.
Front. Artif. Intell. 8:1426455.
doi: 10.3389/frai.2025.1426455

Analyzing handwriting legibility through hand kinematics

Vahan Babushkin[1,2], Haneen Alsuradi[1], Muhamed Osman Al-Khalil[3] and Mohamad Eid[1]*

[1]Applied Interactive Multimedia Lab, Engineering Division, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [2]Tandon School of Engineering, New York University, New York, NY, United States, [3]Arabic Studies Program, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

Introduction: Handwriting is a complex skill that requires coordination between human motor system, sensory perception, cognitive processing, memory retrieval, and linguistic proficiency. Various aspects of hand and stylus kinematics can affect the legibility of a handwritten text. Assessing handwriting legibility is challenging due to variations in experts' cultural and academic backgrounds, which introduce subjectivity biases in evaluations.

Methods: In this paper, we utilize a deep-learning model to analyze kinematic features influencing the legibility of handwriting based on temporal convolutional networks (TCN). Fifty subjects are recruited to complete a 26-word paragraph handwriting task, designed to include all possible orthographic combinations of Arabic characters, during which the hand and stylus movements are recorded. A total of 117 different spatiotemporal features are recorded, and the data collected are used to train the model. Shapley values are used to determine the important hand and stylus kinematics features toward evaluating legibility. Three experts are recruited to label the produced text into different legibility scores. Statistical analysis of the top 6 features is conducted to investigate the differences between features associated with high and low legibility scores.

Results: Although the model trained on stylus kinematics features demonstrates relatively high accuracy (around 76%), where the number of legibility classes can vary between 7 and 8 depending on the expert, the addition of hand kinematics features significantly increases the model accuracy by approximately 10%. Explainability analysis revealed that pressure variability, pen slant (altitude, azimuth), and hand speed components are the most prominent for evaluating legibility across the three experts.

Discussion: The model learns meaningful stylus and hand kinematics features associated with the legibility of handwriting. The hand kinematics features are important for accurate assessment of handwriting legibility. The proposed approach can be used in handwriting learning tools for personalized handwriting skill acquisition as well as for pathology detection and rehabilitation.

KEYWORDS

handwriting, deep learning, temporal convolutional networks, sensorimotor learning, machine learning

1 Introduction

Handwriting is a complex sensorimotor skill that requires simultaneous coordination between human visual-perceptual, cognitive, and motor systems (Bonney, 1992). Writers process visual and haptic feedback to coordinate the hand, arm and finger movement in order to produce legible handwriting. Developing legible handwriting is crucial for

children development, and can affect the educational process, academic success, and self-confidence (Chang and Yu, 2013). Furthermore, understanding the factors that influence handwriting legibility is essential for various practical applications. These include designing personalized learning programs to cater to individual needs (Jenkins et al., 2016), identifying and addressing handwriting difficulties to support students' learning (Drotár and Dobeš, 2020; Fancher et al., 2018), and verifying individuals identities through handwriting analysis in security and forensic contexts (Galbally et al., 2007). Thus, it can be observed that the study of handwriting legibility can inform both educational strategies and technological applications, highlighting its broader significance.

Handwriting legibility is a characteristic of the handwritten text that contributes to its readability (Rosenblum et al., 2004). Handwriting legibility often relies on expert evaluation of the produced handwritten sample (van Drempt et al., 2011). Several global scales are used to evaluate the legibility of healthy adults' handwriting, particularly of medical personnel. For instance, a 4-points scale is utilized to classify the legibility of handwritten medical documents as "illegible", "mostly illegible", "mostly legible", and "legible" (Rodríguez-Vera et al., 2002). Expert evaluation is based on analyzing visual features of the handwritten sample such as size, spacing, alignment, slant, and formation (Amundson and Weil, 1996; Feder and Majnemer, 2007; Fancher et al., 2018). Early studies identified five factors characterizing the legibility of handwriting, namely letter formation, spacing, alignment slant and quality of line (Freeman, 1915). Subsequent studies also suggest that letter formation, size, text alignment and spacing significantly influence the legibility of children's handwriting (Ziviani and Elkins, 1984; Graham et al., 2006). In a recent study, different machine learning approaches were utilized to evaluate the legibility and aesthetics of handwritten text from images of Bengali handwritten documents, reporting 85.74% and 86.69% F-score, for legibility and aesthetics evaluation, respectively (Adak et al., 2017).

Given handwriting is a dynamic process including kinematic, spatial, and temporal components; more objective and quantitative methods are developed based on these dynamic features (Rosenblum et al., 2003). A few studies used the stylus kinematics data and machine learning to evaluate the legibility of individual's signature. For instance, occidental signatures, that incorporate letters and signs into concatenated text with some flourishing elements, are considered (Galbally et al., 2007). To determine if the individual's name can be inferred from the signature, i.e., if signature is legible or illegible, five stylus kinematics features are recorded, including pen-tip coordinates, pressure, and slant, and utilized to engineer 20 global features characterizing the individual's signature. These features are used to train a Multilayer Perceptron (MLP) classifier, achieving 84.56 % accuracy for binary classification (Galbally et al., 2007). Other temporal features are also considered to evaluate the legibility of handwriting, including handwriting speed (Graham et al., 1998a), handwriting style (Graham et al., 1998b), the applied pressure (Harris and Rarick, 1959), and the grasping style (Schwellnus et al., 2012). The potential of these features to detect pathologies such as Alzheimer's disease (AD) from handwriting is explored in Wang et al. (2019), demonstrating that AD patients produce lower pen pressure and variations in the vertical direction, in comparison to healthy subjects.

While hand motion parameters such as fingers/palm position/orientation, acceleration/deceleration, and overall hand speed may influence legibility, they have yet to be explored. This study builds on the methods and findings outlined in PhD dissertation (Babushkin, 2024), and aims to examine correlates between hand/stylus kinematics and handwriting legibility. An experimental setup is developed to complete a handwriting task using a handwriting tablet while recording the stylus and the hand movement in 3D. A deep learning model, inspired by temporal convolutional networks (TCN), is constructed to evaluate the legibility of handwriting based on the time series kinematic data. We hypothesize that hand kinematic features play a prominent role in evaluating handwriting legibility. The interpretable machine learning approach (Shapley values) is used to identify prominent sensorimotor features derived from hand and stylus kinematics in evaluating the handwriting legibility.

2 Methodology

2.1 Experimental setup

The experimental setup (Figure 1) and protocol are based on a previously established methodology (Babushkin et al., 2024, 2023), with modifications to specifically investigate handwriting legibility. The experimental setup (Figure 1A) includes a HUION GT-116 tablet paired with a pen-like stylus and an Ultraleap Stereo IR 170 hand motion tracker. The hand tracking device, as shown in Figure 1A is attached to a rigid stand in a way that allows it to accurately track the writer's hand movements. The system can be easily moved, allowing for data collection in different locations.

2.2 Experimental task and protocol

Similarly to Babushkin et al. (2024), participants were instructed to write a text consisting of 26 Arabic words (Figure 1B), as this number was optimal for fitting within the dimensions of the recording tablet screen. The text was meticulously designed to cover all 28 letters of the Arabic alphabet and to include all key connectivity positions in Arabic orthography. Despite the unique glyphs (e.g., ك) were included in both their connected and unconnected forms, homoglyphs were not individually represented in all their connected/unconnected variations, but rather as a group (for example the ب, ت, ث homoglyphs).

The choice of Arabic script is justified by its cursive nature, context sensitivity, and multiple writing styles, which makes it more complex in comparison to the Latin script (Naz et al., 2013, 2014; Kacem et al., 2012). All these features of Arabic orthography allow to address a wider variety of handwriting skills.

The subject listened to the entire text sample at a speed of 20 words per minute before the start of the experiment. The experiment started when the participant pressed the "Start" button on the tablet screen to start recording of data (hand

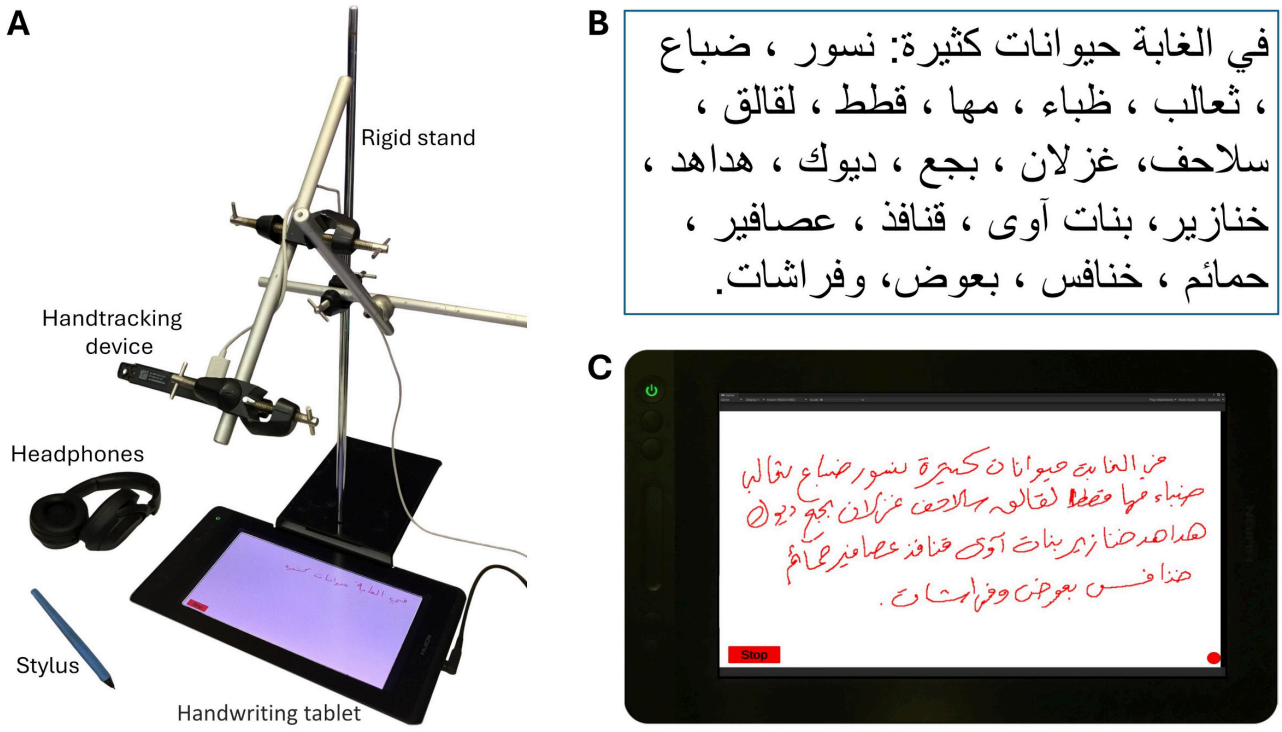


FIGURE 1
(A) Experimental setup, **(B)** sample text dictated to subjects, **(C)** the user interface, displaying a sample of Arabic handwriting recorded from a single subject (Babushkin et al., 2024).

tracking and stylus), which turned red indicating recording was in progress and changed its label to "Stop". The experimenter played the sample text to the subject from an audio recording, adjusting dictation speed according to the pace of the subject's handwriting. As soon as the dictation ended, the subject pressed the "Stop" button to submit the recording. The collected data contained tablet screenshots (Figure 1C), seven stylus-kinematic features recorded from the tablet, and 110 hand-kinematic features recorded by the hand tracking device. The handwriting task was repeated 6 times for each subject with the same text being dictated each time.

2.3 Participants

In total, 50 participants were recruited for this study. All participants were native Arabic speakers, above 18 years, who attended school with Arabic instruction language from grade 1 and with no previous history of neuromuscular or orthopedic dysfunction or dysgraphia. Additional inclusion criteria required participants to be available for in-person sessions to record handwriting tasks and to predominantly use their right hand for writing. The study was conducted in compliance with the Declaration of Helsinki, following its norms and regulations, and with an authorized protocol by the New York University Abu Dhabi Institutional Review Board (IRB: #HRPP-2023-93).

2.4 Expert evaluation and measures

Three Arabic teaching experts (all females, aged 35–55 years) were recruited to evaluate the legibility of the handwriting samples (image-based) using the eligibility evaluation form (see Supplementary Figure S1). To accommodate for the diversity in style of education systems, the experts represented three different educational and cultural backgrounds (Arabic gulf countries, North Africa, and Middle East). Furthermore, experts were recruited based on the following inclusion criteria: (1) having more than 10 years of experience in teaching Arabic handwriting, and (2) currently working in official (statutory work) or extra-official settings (non-statutory work).

The non-language dependent Handwriting Legibility Scale (HLS) (Barnett et al., 2018) was adapted to incorporate features specific to Arabic handwriting, such as aesthetics. The three Arabic teaching experts were tasked with evaluating the handwriting samples in terms of readability (how easy/difficult is it to read this person's handwriting?), space management (was this person able to fit their writing in the space available?), style consistency (how consistent was this person in following specific style?), and aesthetics (how beautiful was the handwriting?). Each of these questions was rated on a 3-point Likert scale, and a cumulative legibility score was calculated by summing the responses to these four questions, yielding a total score between 4 and 12 points. Furthermore, based on the experts' observations of high similarity among samples written by the

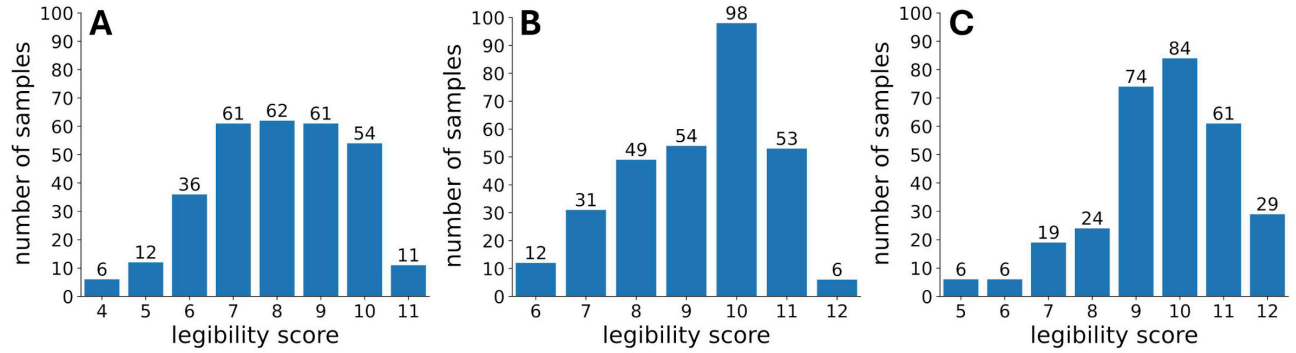


FIGURE 2
Cumulative score distributions by experts **(A)** expert 1, **(B)** expert 2, **(C)** expert 3.

same individual, the cumulative scores of all six samples from each subject were averaged within each subject. The resulting average score was then rounded to the nearest integer and assigned as the legibility score for all samples written by that subject. The cumulative legibility scores assigned by the three experts (see Figure 2) were nearly normally distributed, meaning that the handwriting legibility was well-represented within the recruited subjects.

In total 117 kinematic features were recorded from two different sources: seven stylus features from the tablet and 110 hand kinematics features from the hand tracking device (see Table 1). The data were collected at a sampling rate of 25 Hz and synchronized to a unified timestamp (Babushkin et al., 2024, 2023).

TABLE 1 The recorded 117 hand and stylus kinematics features (Babushkin et al., 2024, 2023).

Modality	Features
Stylus kinematics features	Stylus tip coordinates (x, y, z) ($z = \text{const}$), Pressure (applied force), Azimuth (angle of the stylus projection onto the tablet surface, counted clockwise), Altitude (angle between the tablet screen and the stylus), Proximity to the writing surface.
Hand kinematics features	(x, y, z) coordinates of following Index, Middle, Ring and Pinky fingers' bones: • Distal, • Intermediate, • Proximal, • Metacarpal, • Proximal end of the metacarpal bone. (x, y, z) coordinates of following Thumb bones: • Distal, • Intermediate, • Proximal, • Metacarpal. Palm center (x, y, z) coordinates, Hand pinch position (x, y, z) coordinates (thumb and index if they are pinched), Hand predicted pinch position (x, y, z) coordinates, Hand wrist position (x, y, z) coordinates, Elbow position (x, y, z) coordinates (estimated if not in view), Hand arm center (x, y, z) coordinates (midpoint of the bone), Palm speed (v_x, v_y, v_z) components, Hand palm normal (n_x, n_y, n_z) coordinates, Hand rotational components (r_x, r_y, r_z, r_w), Palm width, Palm pitch, Palm yaw, Palm roll, Hand pinch strength, Hand pinch distance, Hand grab angle, Hand arm length (length of the bone), Hand arm width (average width of flesh around the bone).

2.5 Model architecture

The proposed model is motivated by the temporal convolutional network (TCN) (Lea et al., 2016) design that uses 1D convolutions to extract features encoded across time (Dai et al., 2019). Due to their ability to update layers' weights at every time step simultaneously, TCNs demonstrate better performance than Long-Short Term Memory networks while dealing with long time series (Zhang et al., 2017; Dai et al., 2019); they can take sequences of any lengths and ensure the absence of information leakage from future to past events (Yan et al., 2020). However, TCNs still face difficulties with inferring dependencies between long-range patterns due to the limited receptive field of the convolutional kernels (Dai et al., 2019). The addition of a self-attention layer to TCN enhances its ability to capture these long-range dependencies (Vaswani et al., 2017). Furthermore, the self-attention mechanism allows to infer hidden associations in features, enabling the network to learn irregular and complex patterns (Bu and Cho, 2020). Additionally, self-attention can also lead to more interpretable models (Vaswani et al., 2017).

The proposed model, illustrated in Figure 3, consists of two TCN layers represented by one-dimensional convolutional layers (1D CNN). These are followed by a self-attention layer that processes the hidden representation and extracts a global temporal attention mask. Learning takes place within the subsequent four fully-connected layers. Both the number of fully-connected layers and the number of neurons in each layer were determined empirically, starting with the simplest possible architecture.

The number of TCN layers used in the model is based on the assumption that the first layer captures temporal dependencies, while the second layer focuses on inferring spatial dependencies from the feature maps produced by the first layer. Visualization of feature maps after the first and second 1D CNN layers confirmed this assumption. Adding more than two 1D CNN layers leads to incremental improvements in the model's performance.

The input layer takes a matrix of 117 features and t time points and feeds it to the first convolution layer with number of channels $C_1 = 128$. The original sample of length $T = (w - s) \times n + s$, where s is overlap, is split into n windows of length w. The convolution is performed by sliding a kernel of size $K_1 = 20$ along the time dimension of each window. The resulting $(w - K_1) + 1 \times C_1$ matrix is passed to the input of the second 1D convolution layer of 32 channels with a kernel of size $K_2 = 20$ sliding along the time dimension. The $w - (K_1 + K_2) + 2 \times C_2$ feature map matrix from the second 1D convolution layer is processed by self-attention layer of 32 units, then flattened and passed to the fully connected layers with 256, 128, 64, 32 and finally N neurons, where N is the number of classes. To stabilize the learning process and to prevent overfitting, the batch normalization and dropout of 25% were applied after convolutional and fully connected layers. The Rectified Linear Unit (ReLU) activation was used in all layers except the last one (output layer), which uses Softmax as an activation function. The model was trained on 200 epochs using categorical crossentropy loss function and Adaptive Moment Estimation (Adam) optimizer (Kingma and Ba, 2015). Due to the limited sample size, adjusting the learning rate using callbacks was not feasible – the optimal learning rate of 10^{-3} was found empirically and remained constant during the training. To combat the class imbalance, the oversampling method is applied to the training data, before using it to train the network.

2.6 Feature selection: Shapley values

Shapley values, initially introduced within game theory (Shapley, 1953), are used for assessing the influence of each feature on the prediction of the model (Lundberg and Lee, 2017; Fryer et al., 2021). The core concept behind Shapley values involves assessing the impact of each feature on the model's outcome by sequentially substituting each feature with uniformly-distributed random values and retraining the model on this new dataset. Shapley values are computed by comparing the model's predictions on the dataset with the random feature against the ones on the dataset with the original feature, for all instances in the validation set. These values are then averaged across the validation set to determine the overall influence of each feature. Ultimately, the distribution of averaged Shapley values is obtained, allowing to test the importance of the replaced feature compared to the substituted random feature.

The Shapley value of a feature of index $f \in F = \{1, \ldots, d\}$ from the set of all feature indices F, is a weighted average of all marginal contributions $M_f(S)$, each of them represents the difference in evaluation after introducing feature of index f to a sub-model $S \subset F$, i.e. $M_f(S) = C\left(S \cup \{f\}\right) - C(S)$ (C is evaluation function). In this case, the Shapley value, ϕ_f, of feature, f, is:

$$\phi_f = \sum_{S \in 2^{F \setminus \{f\}}} \omega(S) M_f(S), \tag{1}$$

where $\omega(S) = \frac{|S|!(|F|-|S|-1)!}{|F|}!$ are the weights (Fryer et al., 2021; Babushkin et al., 2024, 2023).

2.7 Statistical analysis

Statistical analysis was conducted to understand how the most prominent features, extracted through Shapley values analysis, differed for samples with high and low legibility. The low and high legibility classes were selected for each expert following the legibility score distributions shown in Figure 2, i.e. for expert 1 the lowest legibility score was 4, highest was 11, for expert 2 lowest was 6, highest was 12, for expert 3 lowest was 5, highest was 12. The top 6 features, consistent across experts, namely pressure, azimuth, altitude and hand speed x, y and z components, were averaged over time for both the lowest and highest legibility classes. To ensure the independence of time-averaged features, the samples evaluated as high or low by more than one expert, were considered once only, i.e. there were no repetitions. The sample size for low legibility group for each feature was 619, and 1,384 for high legibility group. For each feature, the D'Agostino-Pearson omnibus normality test (recommended for large sample sizes D'Agostino and Stephens, 1986) was used to determine if the data follows normal distribution. The results demonstrated that the distributions for low and high legibility groups for all the 6 features were not normal and thus non-parametric Mann-Whitney U test was applied to evaluate statistical difference.

3 Results

3.1 Optimal parameter search

Legible handwriting is commonly evaluated based on the whole handwriting sample rather than a letter or a word (van Drempt et al., 2011). However, due to the limited sample size (303 paragraphs produced by 50 participants) and imbalanced distribution of the samples across the legibility classes, the sliding window method is adopted to enhance the sample size for model training and evaluation. Assuming the legibility of handwriting can vary within the text, the window should be large enough to contain sufficient number of words to ensure sufficient representation of the overall legibility score of the sample.

To determine the optimal length of the time window, the grid-search technique was conducted for 27 different window lengths. Initially, the optimal overlap size was determined by training the model using three different fixed-length windows. Since the length of the window corresponding to the shortest sample was 1,774, the length of largest of these 3 fixed windows was selected as 1,728—a multiple of 64 which was close to 1,774. The smallest window length was set to 64 and the median window length was selected as 896. The overlaps were ranging from 0% to 90% with 10% step. While the number of words per given window of length w varied from subject to subject, the choice of step 64 was dictated by the minimum possible word length. The average accuracy was calculated over 5 folds and 5 runs for each expert and window length. The results indicated a linear increase in accuracy with the percentage of overlap (see Figure 4), leading to the adoption of a 90% overlap.

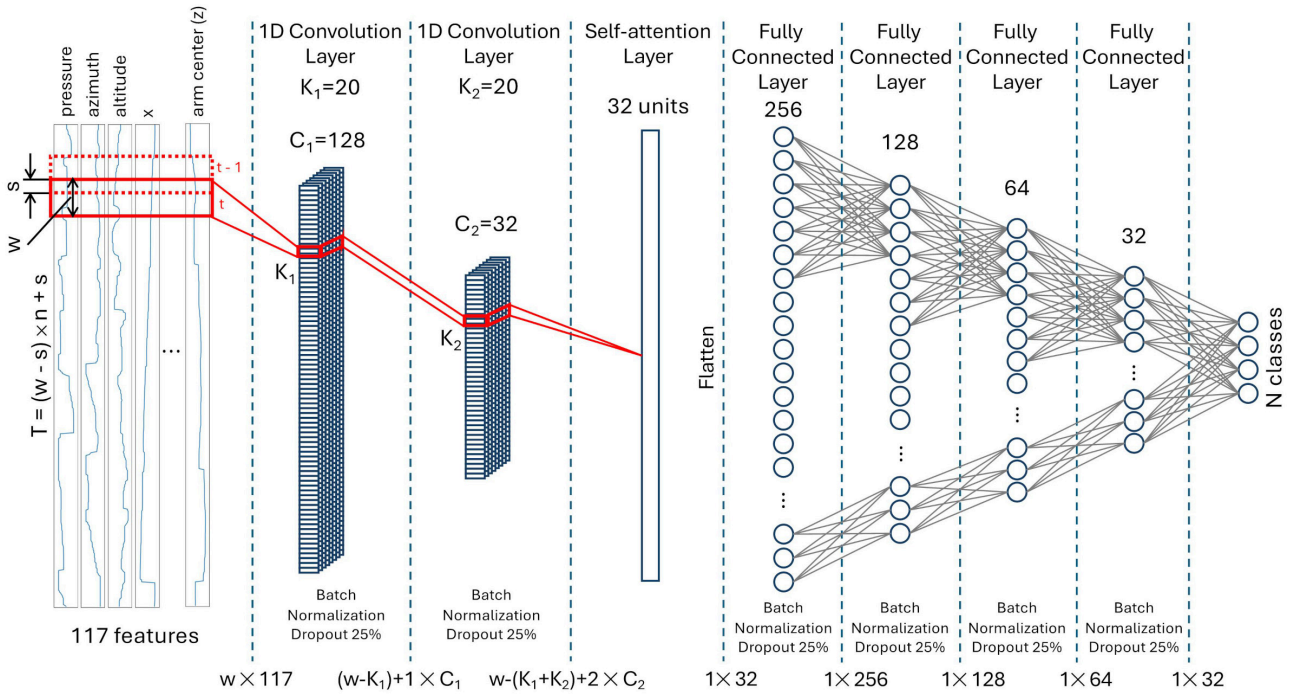


FIGURE 3
The model architecture with two temporal convolution layers. T, is the length of entire paragraph (can vary depending on the writing speed of the subject); n, number of windows; w, length of the window; s, overlap; K_1, K_2, kernel sizes of first/second 1D CNN layers; C_1, C_2, number of channels in first/second 1D CNN layers; N, number of classes.

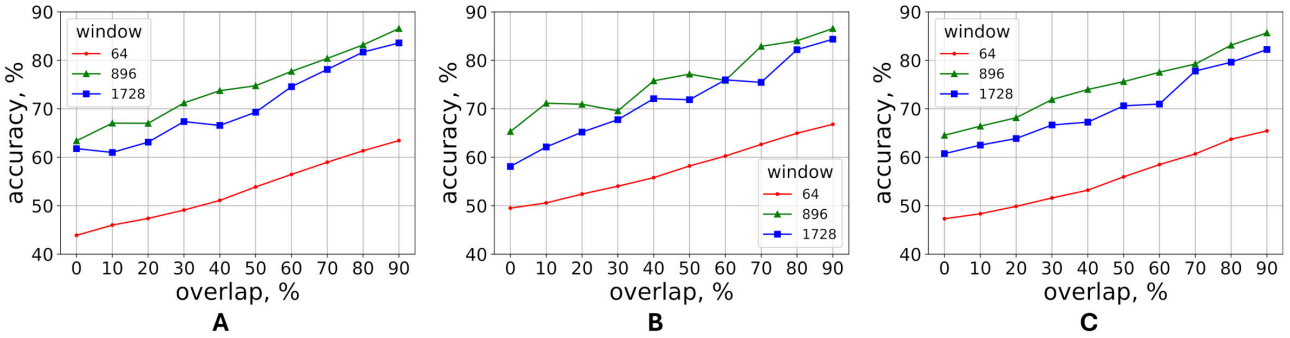


FIGURE 4
Parameters search—optimal overlap. For each expert, the model with optimal hyperparameters was evaluated for three different window sizes using 5-fold cross-validation across 10 evenly spaced overlap percentages values. The process was repeated five times, each with a different random seed. The accuracy values, averaged across five folds and five runs, for **(A)** expert 1, **(B)** expert 2, **(C)** expert 3.

The parameter search was conducted to select a window containing sufficient number of words to justify the assignment of legibility score of the text sample to this window. The optimal window length was estimated by iterating over windows of lengths from 64 to 1728 with the step of 64 and overlap of 90%. For each iteration, the model was trained and validated for 5 folds. The process was repeated 5 times for each expert, each time with different random seed and the accuracy was averaged over all folds for each run. The results, presented in Figure 5, indicated that the model accuracy for each expert increased steadily from a window length of 64 to 576, plateauing around a window size of 1,408, and then declined slowly for larger windows. At window size of 704 all three experts were close to each other reporting an accuracy of around 84%. Therefore, a window size of 704 offered an optimal balance, achieving high accuracy while maintaining a minimized window length, which allowed for a larger number of samples.

3.2 Model performance evaluation

The proposed model was evaluated using 5-fold cross-validation in terms of accuracy, precision, recall and F1-score both with stylus alone and stylus and hand kinematics features. To avoid data leakage,

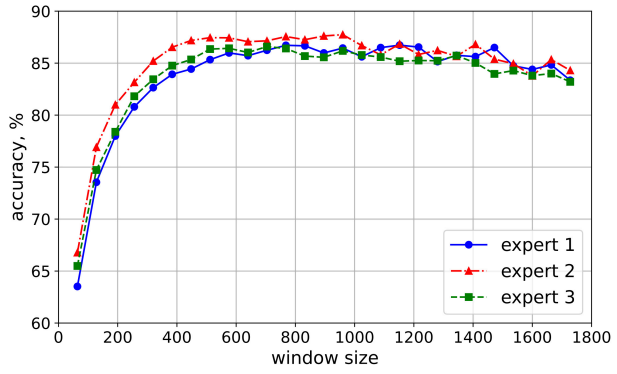


FIGURE 5
Parameters search—optimal window size. For each expert, the model with optimal hyperparameters was evaluated with 5-fold cross-validation across different window sizes, using 90% overlap. The evaluation was repeated five times, each with a different random seed. The accuracy values were averaged across five folds and five runs.

the folds were formed at the paragraph level, before splitting a paragraph with sliding window of length 704 and 90% overlap. Thus all windows from the same paragraph were found only either in 4 folds for training or in the 5-th fold for testing, but not both in the training and testing folds simultaneously. The average performance of the model for each expert was summarized in Table 2. The confusion matrices are demonstrated in Supplementary Figure S2.

Table 2 clearly demonstrates that the proposed model performed consistently well across all experts, with accuracy exceeding 85%. However, the model trained on expert 1 labels achieved the highest F1 score. It might be attributed to the distribution of expert 1 labels across classes being more uniformly distributed. On the other hand, expert 2 and expert 3 distributions were biased toward the high legibility scores, which suggests some leniency in evaluation. This leniency might be attributed to the educational system and cultural background of the last two experts.

3.3 Feature analysis

Shapley values were used to estimate the contribution of each feature in determining the legibility of handwriting. The three models, each trained for one expert's scores with optimal parameters were cross-validated for 5 folds. The Shapley values were evaluated for each fold with 500 samples from the testing set of the given fold. It was recommended to use the testing set to calculate Shapley values to better inspect the ML model and understand the model's decision-making process. However, Shapley values for the testing set allowed evaluating the features impact on the model's generalization performance. For each sample, Shapley values were calculated solely for the class corresponding to the true label of that sample. The obtained Shapley absolute values were averaged across 704 time-points and 500 test instances, and then aggregated for all 5 folds. The 12 most prominent features for predicting the handwriting legibility score by each expert are shown in Figure 6. It is clear that the pressure, hand speed components, and pen slant (altitude, azimuth) are consistently the top features across the three experts.

Figure 7 shows the results of the time-average of the top 4 features for low and high legibility. The average altitude is significantly higher for low legibility than high legibility ($p < 0.01$, Mann-Whitney U Test). On the contrary, azimuth, and hand absolute velocity $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$, where v_x, v_y and v_z are hand speed components, are significantly higher for high legibility as compared to low legibility ($p < 0.01$, Mann-Whitney U Test). According to Harris and Rarick (1959), the pressure itself does not necessarily correlate with legibility for healthy adults, but the pressure variability does. The pressure variability was also calculated as the standard deviation of pressure over time for each time window for high and low legibility classes. The Mann-Whitney U Test confirms that the pressure variability is significantly higher for low legibility ($p < 0.01$), which finds echo in previous literature (Harris and Rarick, 1959). Apparently, the good performance of the model is due to its ability to capture the differences between features from samples coming from different legibility groups.

Finally the correlation between stylus and hand kinematics features such as pressure variability and the absolute speed of handwriting was considered for writers whose handwriting samples (paragraphs) were evaluated as either highly legible or low legible. The paragraphs, produced by each one of those writers were unanimously evaluated either as highly legible or low legible by at least one expert. Highly legible paragraphs corresponded to the highest cumulative legibility scores assigned by the experts (11 for expert 1, 12 for experts 2 and 3; see Figure 2), while low legibility paragraphs corresponded to the lowest scores (4 for expert 1, 6 for expert 2, and 5 for expert 3; see Figure 2). The absolute velocities and pressure values were aggregated across all paragraphs for each subject. The mean absolute velocity and pressure variability were calculated from the aggregated data, resulting in 50 mean absolute velocity—pressure variability pairs, one for each subject. Both features were normalized to $[0, 1]$ interval. The correlation coefficient between mean absolute velocity and pressure variability was 0.19, indicating a very weak relationship between these two features.

4 Discussion

The comparison of model performances with and without hand kinematics features (shown in Table 2) revealed the importance of including hand kinematics for more accurate evaluation of handwriting legibility. While the stylus kinematics features such as applied force, azimuth, and altitude might be sufficient for general legibility assessment, the inclusion of hand kinematics might detect subtle changes in legibility that can be used for diagnosing or predicting handwriting difficulties. Hand kinematics features, measured directly from hand tracking, capture hand dynamics more accurately than tablet features. While handwriting speed can be approximated from stylus kinematics features, it does not explicitly measure the hand speed, which is captured by a hand tracking device. Thus, the inclusion of the hand kinematics features provides the model with a more accurate analysis of the hand dynamics.

The increase in model performance with window overlap for each expert (Figure 4) can be viewed as a form of data augmentation.

TABLE 2 Results of five-fold cross-validation for models using inputs from seven stylus kinematics features and 117 stylus and hand kinematics features for each expert.

	Accuracy, %		Precision, %		Recall, %		F1-score, %	
	Stylus	Stylus and hand	Stylus	Stylus and hand	Stylus	Stylus and hand	Stylus	Stylus and hand
Expert 1	76.0	86.3	78.8	88.4	74.8	89.3	75.8	88.6
Expert 2	76.5	87.2	76.4	86.9	73.2	87.0	73.7	86.4
Expert 3	75.3	85.4	75.8	88.4	73.8	84.5	73.4	85.8

The performance metrics were averaged over five folds.

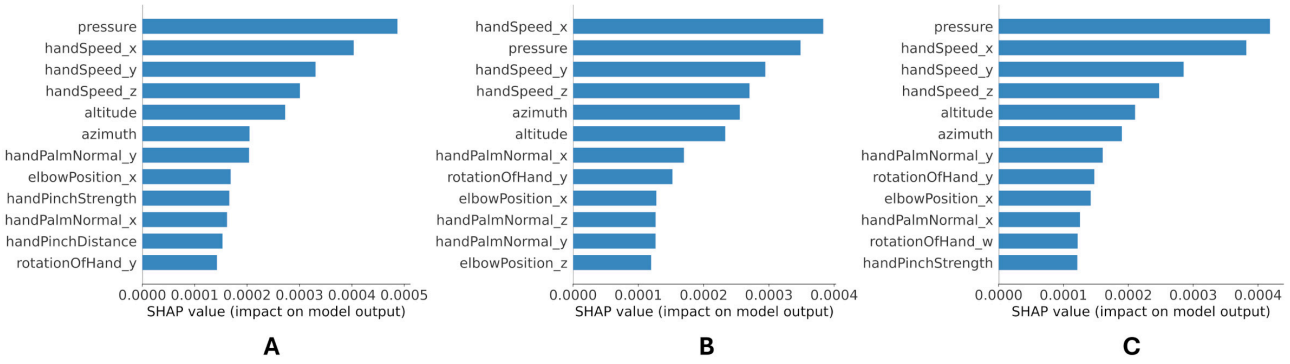


FIGURE 6
Aggregated Shapley values for each expert: **(A)** expert 1, **(B)** expert 2, **(C)** expert 3.

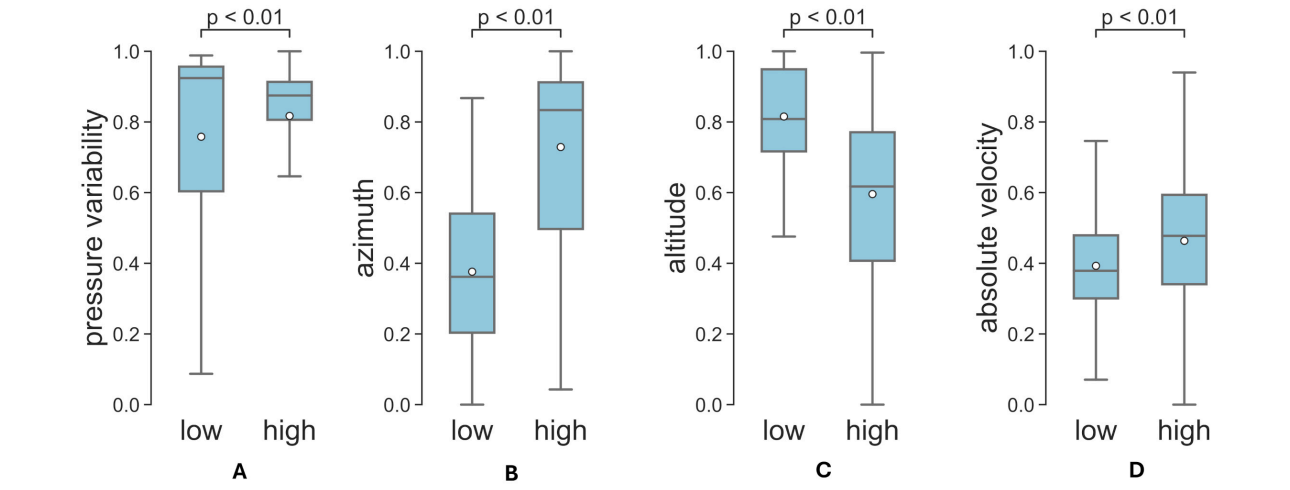


FIGURE 7
Top influential factors for low and high legibility, **(A)** pressure variability, **(B)** azimuth, **(C)** altitude, **(D)** hand absolute velocity. The scales were normalized to $[0, 1]$ interval.

Given the limited number of samples, increasing the overlap percentage between consecutive windows boosts the training dataset size, thereby enhancing accuracy. Moreover, hand and stylus kinematics data are highly temporal, requiring consecutive samples for effective learning. Training with higher overlap enables the model to learn temporal dependencies. Additionally, the data augmentation helps mitigate overfitting by exposing the model to samples with slight variations in hand and stylus kinematics, ensuring better generalization across subjects.

The search for optimal window length revealed another interesting behavior of the model. Despite the fact that shorter windows provide larger number of samples, extremely short window lengths, that contain few to a fragment of a word, are not sufficient to make inference about handwriting legibility (see Figure 5) as smaller textual content within each window may not provide enough information for the model to accurately assess legibility. Consistent force patterns, which are crucial for assessing legibility, are observed at the sentence level rather than the word level (Harris and

Rarick, 1957). Furthermore, since there is an established correlation between the variability of pressure and handwriting legibility (Harris and Rarick, 1959), the model apparently infers the variability of pressure from a longer time interval, equivalent to more than a few words. Similarly, for a human expert it might be hard to provide an accurate evaluation of the legibility of handwriting just by observing a few words. The slight drop in accuracy for larger windows can be explained by the decrease in training samples with the increase of window length as well as the drop in the ability of model to generalize given the increased information per window.

There are significant differences in hand and stylus kinematics features between low and high legibility. Specifically, the pressure variability (Figure 7A) is significantly higher for low legibility than for high legibility, which is also established in previous studies, despite the different approaches to measure the applied force (Harris and Rarick, 1959). Lower overall hand speed (Figure 7D) appears to be associated with low legibility. While previous studies only hinted that subjects who write faster receive lower legibility scores (Harris and Rarick, 1959), our analysis revealed the opposite effect. This finding can be influenced by the differences in how the hand speed is measured—in our study the whole hand is tracked, while in Harris and Rarick (1957) and Harris and Rarick (1959), the hand speed was inferred from the oscillographic records. Other factors that may have influenced this result include the use of electronic tablet (rather than a physical paper) and the lack of friction feedback on the tablet. Interestingly, our study found a very weak correlation between pressure variability and absolute handwriting velocity. This may be due to the fact that pen pressure and writing speed are governed by distinct motor control processes. In adults, pen pressure has been shown to positively correlate with activity in the wrist extensor and flexor muscles, whereas increased writing speed is associated with decreased activation of these distal muscle groups (Gerth and Festman, 2023; Naider-Steinhart and Katz-Leurer, 2007). Additionally, the weak correlation could be influenced by factors such as the non-linear relationship between velocity and pressure variability, as well as individual differences in writing styles among participants. Beyond speed and pressure variability, pen slant (altitude, azimuth) is also a significant factor influencing evaluation of handwriting legibility. Specifically, a smaller azimuth angle and larger altitude features are associated with low legibility. This finds echo in literature where a previous study found that the pen slant is associated with handwriting difficulty (Asselborn et al., 2020).

Despite the differences in experts' cultural and academic backgrounds, the explainability analysis conducted with Shapley values (Figure 6) suggests that the pressure variability, hand speed components, and stylus slant (altitude, azimuth) features are consistently important across experts. This means that in general, experts implicitly rely on these features to evaluate the handwriting legibility. Other features that vary across the experts, are expert-specific and signify the assessment style of each expert. The model explainability analysis can be used to identify features that are correlated with low legibility and suggest handwriting practices/interventions to target these features and improve more effectively the legibility of handwriting.

A few limitations should be acknowledged. First, the sample size of 50 subjects is considered relatively low. The majority of subjects (30) were aged between 18 and 25 years with males constituting only half of the female population. The analysis of kinematic features influencing the legibility of handwriting should be performed on a larger and more diverse sample size. Furthermore, the effect of some demographic parameters, such as age and gender, shall be investigated. In future studies more experts will be recruited to evaluate legibility of the handwriting samples for understanding how the cultural and educational background influences the human-based evaluation and using this knowledge to mitigate possible expert-related biases in cumulative legibility score. Another technology related limitation involves the intermittent hand tracking for some subjects, particularly for female participants mostly due to the hand shape and in some cases the use of sunscreen cream that caused difficulties for tracking with infrared camera. Future research should consider using other data acquisition systems that allow writing on a physical paper (e.g., Wacom Bamboo) or emulating a paper interaction (e.g., reMarkable).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by New York University Abu Dhabi Institutional Review Board (IRB: #HRPP-2020–12). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

VB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. HA: Conceptualization, Investigation, Supervision, Validation, Writing – original draft, Writing – review & editing. MA-K: Validation, Writing – original draft, Writing – review & editing. ME: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This research was supported by the ASPIRE Award for Research Excellence (AARE-2019) program (project number: AARE19-159). This work was also supported by the NYUAD Center for Artificial Intelligence and Robotics (CAIR), funded by Tamkeen under the NYUAD Research Institute Award, CG010.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frai.2025.1426455/full#supplementary-material

References

Adak, C., Chaudhuri, B. B., and Blumenstein, M. (2017). "Legibility and aesthetic analysis of handwriting," in *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* (Kyoto: IEEE), 175–182.

Amundson, S. J., and Weil, M. (1996). Prewriting and handwriting skills. *Occup. Ther. Child.* 3, 524–541.

Asselborn, T., Chapatte, M., and Dillenbourg, P. (2020). Extending the spectrum of dysgraphia: A data driven strategy to estimate handwriting quality. *Sci. Rep.* 10:3140. doi: 10.1038/s41598-020-60011-8

Babushkin, V. (2024). *Analyzing Arabic handwriting through hand* kinematics: A deep learning approach. Brooklyn, New York: New York University Tandon School of Engineering.

Babushkin, V., Alsuradi, H., Al-Khalil, M. O., and Eid, M. (2024). Analyzing Arabic handwriting style through hand kinematics. *Sensors* 24:6357. doi: 10.3390/s24196357

Babushkin, V., Alsuradi, H., Jamil, M. H., Al-Khalil, M. O., and Eid, M. (2023). Assessing handwriting task difficulty levels through kinematic features: a deep-learning approach. *Front. Robot. AI* 10:1193388. doi: 10.3389/frobt.2023.1193388

Barnett, A. L., Prunty, M., and Rosenblum, S. (2018). Development of the handwriting legibility scale (HLS): A preliminary examination of reliability and validity. *Res. Dev. Disabil.* 72, 240–247. doi: 10.1016/j.ridd.2017.11.013

Bonney, M.-A. (1992). Understanding and assessing handwriting difficulty: perspectives from the literature. *Aust. Occup. Ther. J.* 39, 7–15. doi: 10.1111/j.1440-1630.1992.tb01751.x

Bu, S.-J., and Cho, S.-B. (2020). Time series forecasting with multi-headed attention-based deep learning for residential energy consumption. *Energies* 13:4722. doi: 10.3390/en13184722

Chang, S.-H., and Yu, N.-Y. (2013). Handwriting movement analyses comparing first and second graders with normal or dysgraphic characteristics. *Res. Dev. Disabil.* 34, 2433–2441. doi: 10.1016/j.ridd.2013.02.028

D'Agostino, R. B., and Stephens, M. (1986). *Tests for Normal Distribution in Goodness-of-fit Techniques*. New York: Marcel Decker.

Dai, R., Minciullo, L., Garattoni, L., Francesca, G., and Bremond, F. (2019). "Self-attention temporal convolutional network for long-term daily living activity detection," in *2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)* (Taipei: IEEE), 1–7.

Drotár, P., and Dobeš, M. (2020). Dysgraphia detection through machine learning. *Sci. Rep.* 10:21541. doi: 10.1038/s41598-020-78611-9

Fancher, L. A., Priestley-Hopkins, D. A., and Jeffries, L. M. (2018). Handwriting acquisition and intervention: a systematic review. *J. Occup. Ther. SchoolsEarly Intervent.* 11, 454–473. doi: 10.1080/19411243.2018.1534634

Feder, K. P., and Majnemer, A. (2007). Handwriting development, competency, and intervention. *Dev. Med. Child Neurol.* 49, 312–317. doi: 10.1111/j.1469-8749.2007.00312.x

Freeman, F. N. (1915). An analytical scale for judging handwriting. *Elem. Sch. J.* 15, 432–441. doi: 10.1086/454438

Fryer, D., Strümke, I., and Nguyen, H. (2021). Shapley values for feature selection: the good, the bad, and the axioms. *IEEE Access* 9, 144352–144360. doi: 10.1109/ACCESS.2021.3119110

Galbally, J., Fierrez, J., and Ortega-Garcia, J. (2007). "Classification of handwritten signatures based on name legibility," in *Biometric Technology for Human Identification IV* (Bellingham: SPIE), 64–72.

Gerth, S., and Festman, J. (2023). Muscle activity during handwriting on a tablet: an electromyographic analysis of the writing process in children and adults. *Children* 10:748. doi: 10.3390/children10040748

Graham, S., Berninger, V., Weintraub, N., and Schafer, W. (1998a). Development of handwriting speed and legibility in grades 1-9. *J. Educ. Res.* 92, 42–52. doi: 10.1080/00220679809597574

Graham, S., Struck, M., Santoro, J., and Berninger, V. W. (2006). Dimensions of good and poor handwriting legibility in first and second graders: Motor programs, visual-spatial arrangement, and letter formation parameter setting. *Dev. Neuropsychol.* 29, 43–60. doi: 10.1207/s15326942dn2901_4

Graham, S., Weintraub, N., and Berninger, V. W. (1998b). The relationship between handwriting style and speed and legibility. *J. Educ. Res.* 91, 290–297. doi: 10.1080/00220679809597556

Harris, T. L., and Rarick, G. L. (1957). The problem of pressure in handwriting. *J. Exp. Educ.* 26, 151–178. doi: 10.1080/00220973.1957.11010592

Harris, T. L., and Rarick, G. L. (1959). The relationship between handwriting pressure and legibility of handwriting in children and adolescents. *J. Exp. Educ.* 28, 65–84. doi: 10.1080/00220973.1959.11010642

Jenkins, S., Williams, M., Moyer, J., George, M., and Foster, E. (2016). "The shifting paradigm of teaching: Personalized learning according to teachers," in *Knowledge Works*. Retrieved from: https://knowledgeworks.org/wp-content/uploads/2018/01/teacher-conditions.pdf (accessed March 14, 2025).

Kacem, A., Aouïti, N., and Belaïd, A. (2012). "Structural features extraction for handwritten arabic personal names recognition," in *Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition, ICFHR'12* (IEEE Computer Society), 268–273. doi: 10.1109/ICFHR.2012.276

Kingma, D., and Ba, J. (2015). "Adam: A method for stochastic optimization," in *International Conference on Learning Representations (ICLR)* (San Diega, CA: ICLR).

Lea, C., Vidal, R., Reiter, A., and Hager, G. D. (2016). "Temporal convolutional networks: A unified approach to action segmentation," in *Computer Vision-ECCV 2016 Workshops*, eds. G. Hua, and H. Jégou (Cham: Springer International Publishing), 47–54.

Lundberg, S. M., and Lee, S.-I. (2017). "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30*, eds. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et al. (New York: Curran Associates, Inc), 4765–4774.

Naider-Steinhart, S., and Katz-Leurer, M. (2007). Analysis of proximal and distal muscle activity during handwriting tasks. *Am. J. Occup. Ther.* 61, 392–398. doi: 10.5014/ajot.61.4.392

Naz, S., Hayat, K., Razzak, M. I., Anwar, M. W., and Akbar, H. (2013). "Arabic script based language character recognition: Nasta'liq vs naskh analysis", in 2013 *World Congress on Computer and Information Technology (WCCIT)*, 1–7. doi: 10.1109/WCCIT.2013.6618740

Naz, S., Razzak, M. I., Hayat, K., Anwar, M. W., and Khan, S. Z. (2014). "Challenges in baseline detection of arabic script based languages," in *Intelligent Systems for Science and Information: Extended and Selected Results from the Science and Information Conference 2013* (Cham: Springer), 181–196.

Rodríguez-Vera, F. J., Marin, Y., Sanchez, A., Borrachero, C., and Pujol, E. (2002). Illegible handwriting in medical records. *J. R. Soc. Med.* 95, 545–546. doi: 10.1258/jrsm.95.11.545

Rosenblum, S., Parush, S., and Weiss, P. L. (2003). Computerized temporal handwriting characteristics of proficient and non-proficient handwriters. *The American Journal of Occup. Therapy* 57, 129–138. doi: 10.5014/ajot.57.2.129

Rosenblum, S., Weiss, P. L., and Parush, S. (2004). Handwriting evaluation for developmental dysgraphia: process versus product. *Read. Writ.* 17, 433–458. doi: 10.1023/B:READ.0000044596.91833.55

Schwellnus, H., Carnahan, H., Kushki, A., Polatajko, H., Missiuna, C., and Chau, T. (2012). Effect of pencil grasp on the speed and legibility of handwriting in children. *Am. J. Occup. Therapy* 66, 718–726. doi: 10.5014/ajot.2012.004515

Shapley, L. (1953). A value for n-persons games. *Ann. Mathem. Stud.* 28, 307–318. doi: 10.1515/9781400881970-018

van Drempt, N., McCluskey, A., and Lannin, N. A. (2011). A review of factors that influence adult handwriting performance. *Aust. Occup. Ther. J.* 58, 321–328. doi: 10.1111/j.1440-1630.2011.00960.x

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). "Attention is all you need," in *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17* (Red Hook, NY: Curran Associates Inc.), 6000–6010.

Wang, Z., Abazid, M., Houmani, N., Garcia-Salicetti, S., and Rigaud, A.-S. (2019). Online signature analysis for characterizing early stage Alzheimer's disease: a feasibility study. *Entropy* 21:e21100956. doi: 10.3390/e21100956

Yan, J., Mu, L., Wang, L., Ranjan, R., and Zomaya, A. (2020). Temporal convolutional networks for the advance prediction of ENSO. *Sci. Rep.* 10:8055. doi: 10.1038/s41598-020-65070-5

Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L. (2017). Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *Trans. Img. Proc.* 26, 3142–3155. doi: 10.1109/TIP.2017.2662206

Ziviani, J., and Elkins, J. (1984). An evaluation of handwriting performance. *Educ. Rev.* 36, 249–261. doi: 10.1080/0013191840360304

frontiers | Frontiers in Artificial Intelligence

TYPE Review
PUBLISHED 26 June 2025
DOI 10.3389/frai.2025.1585629

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Jie Yang,
Chongqing University of Posts and Telecommunications, China
Praveen Joe I. R.,
Vellore Institute of Technology (VIT), India
Ahmed M. Alaa,
Badr University in Cairo, Egypt

*CORRESPONDENCE
Jing Du
✉ eric.du@essie.ufl.edu

RECEIVED 01 March 2025
ACCEPTED 13 June 2025
PUBLISHED 26 June 2025

CITATION
Wu J, You H and Du J (2025) AI generations: from AI 1.0 to AI 4.0.
Front. Artif. Intell. 8:1585629.
doi: 10.3389/frai.2025.1585629

AI generations: from AI 1.0 to AI 4.0

Jiahao Wu, Hengxu You and Jing Du*

Informatics, Cobots and Intelligent Construction Lab, Engineering School of Sustainable Infrastructure and Environment, University of Florida, Gainesville, FL, United States

This paper proposes that Artificial Intelligence (AI) progresses through several overlapping generations: AI 1.0 (Information AI), AI 2.0 (Agentic AI), AI 3.0 (Physical AI), and a speculative AI 4.0 (Conscious AI). Each AI generation is driven by shifting priorities among algorithms, computing power, and data. AI 1.0 accompanied breakthroughs in pattern recognition and information processing, fueling advances in computer vision, natural language processing, and recommendation systems. AI 2.0 is built on these foundations through real-time decision-making in digital environments, leveraging reinforcement learning and adaptive planning for agentic AI applications. AI 3.0 extended intelligence into physical contexts, integrating robotics, autonomous vehicles, and sensor-fused control systems to act in uncertain real-world settings. Building on these developments, the proposed AI 4.0 puts forward the bold vision of self-directed AI capable of setting its own goals, orchestrating complex training regimens, and possibly exhibiting elements of machine consciousness. This paper traces the historical foundations of AI across roughly 70 years, mapping how changes in technological bottlenecks from algorithmic innovation to high-performance computing to specialized data have stimulated each generational leap. It further highlights the ongoing synergies among AI 1.0, 2.0, 3.0, and 4.0, and explores the ethical, regulatory, and philosophical challenges that arise when artificial systems approach (or aspire to) human-like autonomy. Ultimately, understanding these evolutions and their interdependencies is pivotal for guiding future research, crafting responsible governance, and ensuring that AI's transformative potential benefits society.

KEYWORDS

artificial intelligence evolution, machine learning, reinforcement learning, large language models, AI ethics and governance

1 Introduction

Artificial Intelligence (AI) has experienced a transformative evolution over the last 70 years, evolving from its nascent stage of theoretical formulations to its current status as a cornerstone of technological advancement (Haenlein and Kaplan, 2019). Initially, the field was dominated by intellectual explorations into symbolic reasoning, knowledge representation, and the rudimentary principles of machine learning (Newell and Simon, 1956). These early stages were marked by a focus on conceptual breakthroughs, laying the groundwork for what AI could potentially achieve. As computational capabilities expanded and data sources proliferated, AI transitioned from theoretical models to practical applications capable of learning from patterns and making precise predictions (Alom et al., 2018). The last two decades, however, have witnessed an unprecedented acceleration in AI development, propelling the field into realms that surpass even the most optimistic projections of its early pioneers.

Despite remarkable successes in areas like natural language processing, computer vision, and large-scale data analytics, AI continues to face challenges in interacting seamlessly with

complex, dynamic real-world environments. This ongoing struggle signals an emerging phase in AI's evolution, marking a shift from systems that primarily process and predict information to ones that can plan, decide, and act, ushering in new generations of AI: *Information AI (AI 1.0)*, *Agentic AI (AI 2.0)*, *Physical AI (AI 3.0)* and *Conscious AI (AI 4.0)*. This classification not only clarifies the conceptual transitions within the field but also helps delineate the evolution of AI capabilities from data extraction to making autonomous decisions in digital realms, and now to engaging directly with the physical world.

Understanding these transitions is essential, not just from a technological standpoint but also for grasping the societal and economic implications of AI. Distinct technological drivers and bottlenecks have shaped each phase of AI: the early period was limited by the lack of advanced algorithms and computational frameworks (Jones, 1994); the advent of powerful GPUs around 2012 significantly shifted the landscape, enabling more complex neural architectures (Nvidia, 2011); and today, the challenge has moved toward harnessing domain-specific, high-quality data to feed into these sophisticated systems (Budach et al., 2022). Recognizing these shifts is crucial for stakeholders, including policymakers, researchers, and industry leaders, who must navigate the ethical, regulatory, and technical complexities introduced by advanced AI systems.

This review aims to provide a comprehensive retrospective on the milestones that have defined AI's progress. By tracing the lineage of algorithmic innovations, increases in computing power, and enhancements in data utilization, we aim to illuminate the significant moments that have shaped AI from its inception to its current state. This exploration is structured around the AI 1.0 to AI 4.0 framework, illustrating how each generation's defining features and limitations correspond to broader historical phases from approximately 1950 to the present. In doing so, we will also contemplate the future trajectory of AI, considering the potential technical challenges, societal impacts, and strategic directions that could define the next phases of AI research and application.

This article is structured first to revisit the historical foundations of AI, emphasizing the shifts in primary drivers from algorithms to computing power to data. We then delve into the specific characteristics, achievements, and limitations of AI 1.0, AI 2.0, AI 3.0, and AI 4.0. Following this, we explore AI's convergence and future outlook, highlighting the synergies among the four generations and outlining the grand challenges that lie ahead. Finally, we conclude with a synthesis of key insights and propose future directions for sustained progress in the field, aiming to both inform and inspire continued innovation and thoughtful integration of AI into our daily lives and societal structures.

2 Historical foundations of AI

2.1 Phase 1 (1950s–2010s): Age of algorithmic innovations

Since the 1950s, AI has advanced through a dynamic interplay among three core ingredients: *algorithms*, *computing power*, and *data* (Schmidhuber, 2022). Although these three factors have always shaped the field, they have not always contributed equally at every stage. In the early decades, the limiting factor was innovation in algorithms. From mid-century debates about the feasibility of machine intelligence to the emergence of expert systems and neural networks, it was clear that conceptual breakthroughs would determine AI's boundaries (Turing, 2009). Meanwhile, although data and computing power were important, they played more supportive roles. Gradually, as new hardware architectures appeared and large-scale datasets became more accessible, the focus shifted toward harnessing immense computational capability and vast amounts of information. During this era, most funding for algorithmic research came from government programs (e.g., DARPA's Strategic Computing Initiative) and a handful of industrial labs, fostering tight collaborations between computer scientists, control engineers, and cognitive psychologists to maximize limited hardware through smarter algorithms.

From the outset, researchers were fascinated by whether machines could truly think. Alan Turing's pioneering paper (Turing, 1950) set the stage, posing the famous "*imitation game*" as a litmus test for intelligence. In 1956, the Dartmouth Conference (McCarthy et al., 2006) formally introduced the term "Artificial Intelligence" and laid out the bold proposition that the essence of human intelligence could be precisely described and replicated in machines. Early NSF and DARPA grants enabled interdisciplinary AI centers at MIT and Stanford, where mathematicians, linguists, and early cybernetics experts worked side by side to turn the Turing Test and Dartmouth vision into functioning prototype systems. Early systems, such as the *Logic Theorist* and the *General Problem Solver* (Newell and Simon, 1956; Newell et al., 1959) underscored that symbolic reasoning could be computationally realized. These proof-of-concept attempts highlighted the central premise of that era: if we could devise the right algorithms, computers might reason and solve problems with near-human efficacy.

By the 1960s and 1970s, a strong emphasis on *symbolic AI* emerged. Influential works by McCarthy (1960) introduced LISP as a language suited to symbolic processing, while Minsky and Papert's (1969) critical analysis of single-layer perceptions contributed to a pause in neural network research, pushing many researchers toward knowledge-based or "expert" systems. Milestones like the DENDRAL project (Buchanan et al., 1971) and MYCIN (Shortliffe et al., 1975) showcased how carefully curated rule sets could guide problem-solving in specialized domains. These systems illustrated the power of algorithmic design in areas such as medical diagnosis or chemical analysis, even when real-world data were scarce and computational resources were limited. Corporate and university partnerships in domains like healthcare (e.g., with Stanford Medical School) and chemical analysis funded expert-system projects, creating joint labs where domain specialists and AI researchers codified knowledge bases despite constrained memory and CPU budgets.

Neural networks rebounded in the 1980s with work on Hopfield networks (Hopfield, 1982) (Figure 1) and, crucially, the rediscovery of backpropagation (Rumelhart et al., 1986). This gave researchers fresh insight into how machines might learn patterns from data. Though the potential of these connectionist approaches was clear, they often stalled because large datasets were not widely available and specialized hardware did not yet exist. Even so, foundational contributions like LeCun et al. (1989) application of *convolutional neural networks* to handwritten digit recognition laid the groundwork for modern deep learning. Modest government programs and early industry prototypes, such as the Connection Machine, emerged from collaborations between neuroscientists and computer engineers. Still, widespread

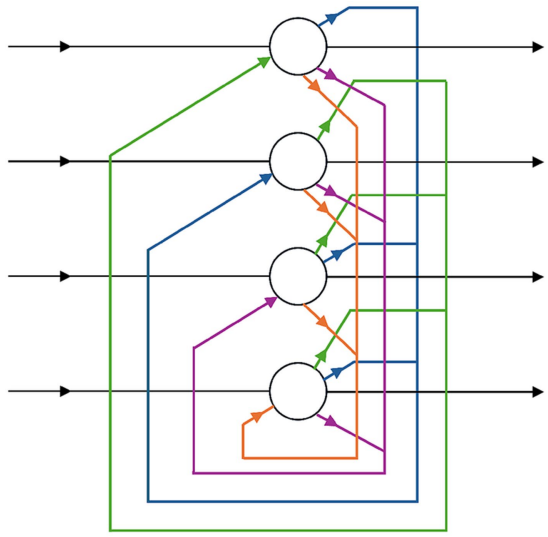


FIGURE 1
The Hopfield networks (Hopfield, 1982) introduced content-addressable memory in neural networks, marking a major milestone in connectionism in AI.

adoption had to await later GPU cost declines and cloud-computing services.

By the 1990s, specific algorithmic achievements hinted at deeper architectures capable of tackling increasingly complex tasks. The proposal of Long Short-Term Memory (LSTM) networks effectively addressed the vanishing gradient problem, opening possibilities for modeling sequential data more accurately (Hochreiter, 1997). However, the real transformative moment emerged around 2012, when Krizhevsky, Sutskever, and Hinton demonstrated that ImageNet-scale datasets and high-performance GPUs could dramatically improve a deep neural network's ability to classify images, i.e., the *AlexNet* (Krizhevsky et al., 2012) (Figure 2). Although this watershed event is often viewed as the dawn of the "deep learning era," it could not have happened without the algorithmic groundwork laid over the preceding decades. The 2012 AlexNet breakthrough itself was propelled by the ImageNet consortium, uniting academic vision labs and industry hardware vendors, and by the sudden availability of affordable GPU clusters donated or subsidized by major tech companies.

2.2 Phase 2 (2010s–present): The computing revolution and deep learning renaissance

The pattern-matching architectures pioneered in AI 1.0, such as convolutional filters for edge and shape detection and Hopfield networks for associative memory, laid the essential groundwork for AI 2.0's learned feature hierarchies. By encoding low and mid-level visual and sequential patterns in trainable layers, these early connectionist models enabled decision-making agents to operate on rich, automatically extracted representations rather than raw sensor data, accelerating reinforcement-learning and supervised learning breakthroughs once sufficient data and compute became available.

A dramatic shift in AI research took hold around 2012, when mounting computational capacity began to eclipse algorithmic novelty as the principal engine of progress. This transition was underwritten by rapidly declining GPU prices, driven by consumer gaming markets, and by major cloud providers (AWS, Google Cloud, Azure) offering GPU instances, which democratized access to parallel computing. Key partnerships between hardware vendors (NVIDIA, AMD) and academic labs established early benchmarks for large-scale training, exemplifying how economic incentives catalyze scientific breakthroughs. While the core concepts underlying neural networks had been present since at least the 1980s, it was the widespread adoption of General-Purpose Graphics Processing Units (GPUs) that ignited what is often termed the "*deep learning renaissance*" (Figure 3) (Nickolls et al., 2008). Collaborative consortia, such as the ImageNet project, brought together vision researchers, software engineers, and data curators from both academia and industry, creating shared data resources and open-source codebases that accelerated innovation and reproducibility. When Krizhevsky et al. (2012) leveraged GPUs to train a large convolutional neural network for the ImageNet competition, they decisively demonstrated how parallelized computing could unearth performance gains previously unachievable with single-threaded Central Processing Units (CPUs). This milestone was enabled by government grants and corporate research labs (e.g., Google Brain, Microsoft Research), which invested in GPU clusters and supported interdisciplinary teams of machine-learning scientists and systems engineers to push the limits of scale. This turning point catalyzed a wave of research across machine vision, speech recognition, and natural language processing, with groups at Google, Microsoft, Baidu, and many academic institutions all racing to scale up network architectures (Dean et al., 2012; Bishop, 2013; Yu, 2013). The essence of this period lay in the conviction that "bigger is better," whether in terms of model parameters, dataset size, or sheer computational resources. Consequently, much of the state-of-the-art progress hinged on harnessing specialized hardware: first GPUs, then tensor processing units (TPUs) and other custom accelerators, to churn through ever-growing datasets in shorter training cycles.

By the mid-2010s, the explosive rise of deep reinforcement learning (Mnih et al., 2015) and breakthroughs in game-playing AI, such as AlphaGo (Silver et al., 2016), underscored that not only could AI models learn representations from massive data, but they could also discover winning strategies through large-scale simulations. These advances were propelled by collaborations between AI theorists, neuroscientists studying decision-making, and high-performance computing experts, as well as by significant venture-capital funding in AI startups focusing on simulation-based learning and autonomous agents. Nevertheless, the predominant realm for these systems remained resolutely digital. Whether classifying images, translating text (Bahdanau, 2014; Vaswani, 2017), or playing complex board and video games, AI was still operating in an essentially informational context. Although data availability was critical and algorithms like convolutional and recurrent neural networks continued to improve, sheer computational power was often the deciding factor in achieving superior performance. Researchers observed emergent patterns in "*scaling laws*" (Kaplan et al., 2020), revealing that larger models trained on larger datasets could unlock qualitatively new capabilities. Systems like GPT-2 (Radford et al., 2019) and GPT-3 (Brown et al., 2020) illustrated this phenomenon vividly by demonstrating a striking ability to generate human-like text once parameter counts and training

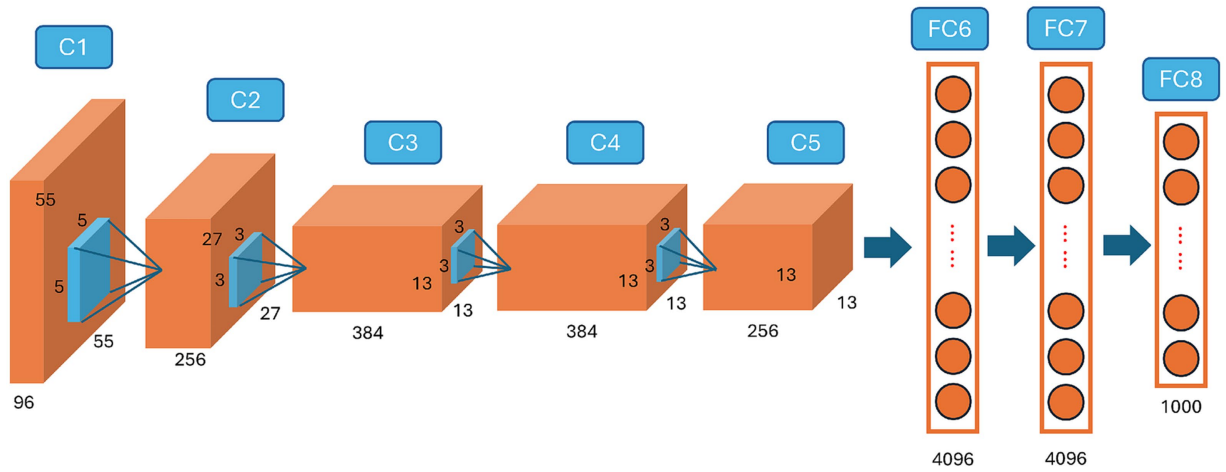


FIGURE 2
AlexNet (Krizhevsky et al., 2012) marks the beginning of large-scale, GPU-accelerated convolutional neural networks for high-performance image classification.

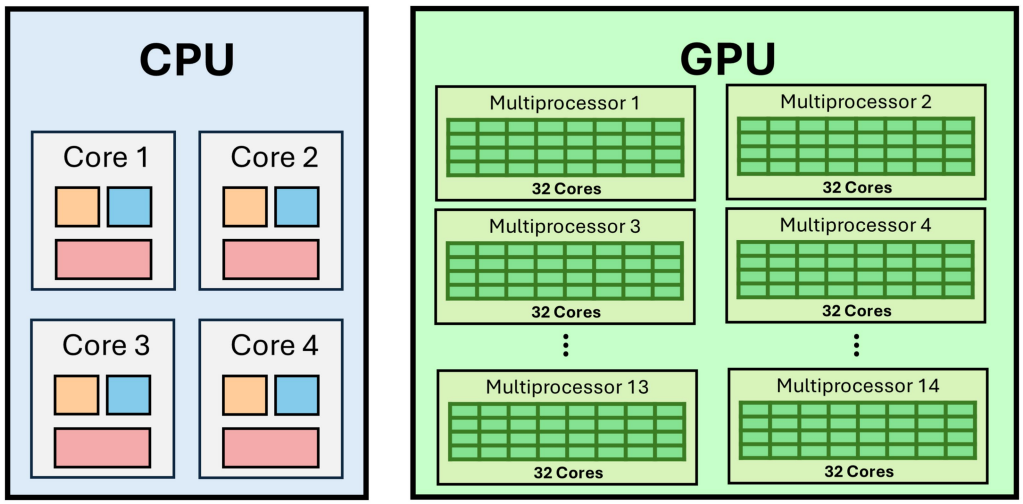


FIGURE 3
The CUDA architecture pioneered general-purpose GPU computing, revolutionizing parallel processing and accelerating AI breakthroughs.

data reached certain thresholds. The development and deployment of these language models were driven by multi-institutional efforts, including OpenAI's partnerships with cloud providers and academic collaborators, and by economic incentives from industries eager to commercialize natural-language interfaces, fueling research consortia around ethical and scalable model training. Because of their sophistication, these models continued to reside in the digital world, making them refined and powerful versions focused on big data analytics and pattern recognition at an unprecedented scale. Even so, the end of this phase began to hint at a transition toward greater autonomy and decision-making in digital contexts, an emerging hallmark of agentic AI. While many systems are still centered on classification or prediction, the rise of advanced reinforcement learning agents able to adapt strategies within software ecosystems foreshadowed a new kind of agency. By approximately 2024, the scholarly and commercial drive to develop goal-directed virtual assistants, automated resource allocation tools, and multi-agent simulations suggested that the chief challenge was no longer purely to label data accurately, but to act in digital environments in ways that transcended traditional supervised learning (Chen et al., 2023). This growing desire for agentic AI remained tied to abundant computing power, yet it began to reveal new dependencies on specialized data streams and real-time feedback loops (Tosi et al., 2024). It set the stage for the next generation of AI, where computational needs would remain vital. Still, data and context-specific knowledge would become even more pivotal in enabling truly autonomous, adaptive systems.

However, this unprecedented shift toward data-driven and compute-driven breakthroughs has also exposed systemic vulnerabilities that must be carefully examined before embracing the next wave of autonomous, agentic AI. The deployment of AI 2.0 into high-stakes domains such as finance, public policy, and healthcare has revealed how tightly coupled, speed-optimized systems can trigger

cascading failures under stress. In financial markets, for instance, high-frequency trading algorithms, tuned to exploit sub-millisecond price discrepancies, precipitated the "Flash Crash" of May 6, 2010, when linked bots erased nearly $1 trillion in equity value within minutes before a partial rebound (Kirilenko et al., 2017). Without unified circuit breakers or oversight mechanisms, these agents amplified feedback loops during extreme volatility, demonstrating that raw performance can come at the price of systemic stability. In law enforcement, predictive-policing tools trained on decades of arrest data in Chicago and Los Angeles disproportionately targeted minority neighborhoods, perpetuating historical biases and eroding community trust. Credit-scoring models have likewise been shown to underprice loans for underrepresented groups, prompting regulatory investigations into discriminatory lending practices. These examples underscore that high accuracy on benchmark datasets does not guarantee equitable or safe outcomes in complex, real-world settings. Healthcare AI offers a further cautionary tale: diagnostic assistants trained on skewed image collections have misclassified critical conditions in underrepresented populations. A prominent 2018 study found that a melanoma detection model, trained predominantly on light-skinned images, misdiagnosed darker-skinned patients at twice the rate of lighter-skinned counterparts, despite reporting >95% accuracy on its test set. Such failures highlight the dangers of blindly scaling models without rigorous data curation and validation protocols. To guard against these risks, we recommend a multilayered defense: comprehensive stress testing under extreme or adversarial conditions; mandated transparency of model architectures and data provenance to facilitate third-party audits; regulatory circuit-breakers and human-in-the-loop overrides in mission-critical systems; and the formation of interdisciplinary oversight bodies that bring together AI practitioners, ethicists, domain experts, and policymakers. Embedding these safeguards will enable the AI community to harness the power of deep learning while preserving social, economic, and ethical stability.

2.3 Phase 3 (2024–foreseeable future): Data-centric paradigms

In the wake of a period defined by dramatic increases in computational horsepower, the focal point of AI advancement has shifted once again. This transition has been propelled by major industry investments, particularly from cloud providers (AWS, Google Cloud) offering specialized GPU/TPU instances, and by venture capital funding in data-centric startups, which drove large-scale data-aggregation platforms and new public-private consortia to curate domain-specific datasets. Where Phase 2 thrived on scaling neural networks through unprecedented parallel processing, Phase 3 acknowledges that data, especially specialized, high-quality data, is frequently the greatest obstacle. Researchers have discovered that ever-larger models alone do not guarantee success without context-rich training sets. Consequently, large-scale, domain-specific data-collection efforts have emerged, reshaping the field's priorities. Projects that aggregate specialized medical data for diagnostic systems (Topol, 2019), simulate high-fidelity environments for robotics and autonomous vehicles (Dosovitskiy et al., 2017; Kalashnikov et al., 2018), or compile deep reinforcement learning benchmarks with realistic constraints (Bellemare et al., 2013; Dulac-Arnold et al., 2021) attest to the idea that harnessing robust datasets can be as determinative as algorithmic ingenuity or raw computational power.

Despite the continued importance of parallel computing and innovative architectures, many cutting-edge successes now hinge on data strategy. The "data-centric AI" movement gained traction through collaborations between academic labs (e.g., Stanford's DAWN project) and industry partners in healthcare, automotive, and finance, where structured data pipelines and synthetic data initiatives (e.g., NVIDIA's DRIVESim) received dedicated research grants and created shared benchmarks. Researchers have championed "data-centric AI" (Ng, 2021), arguing that refining training sets, removing biases, filling in coverage gaps, or generating synthetic data to handle edge cases, often yields more improvement than adding layers to a neural network. This philosophy is closely related to the rise of foundation models (Bommasani et al., 2021), which are vast neural architectures that can be adapted to myriad tasks but require massive, carefully curated corpora to realize their full potential. As data becomes the true bottleneck, teams must grapple with the logistical and ethical challenges of collecting, storing, and labeling it, as well as with privacy, consent, and representation issues.

Within this phase, AI's transition from informational analysis to agentic decision-making becomes increasingly tangible. Interdisciplinary teams combining roboticists, control engineers, and ethicists, backed by government programs like the U. S. National Robotics Initiative and by multinational R&D labs (e.g., Toyota Research Institute), have spearheaded projects in autonomous vehicles, surgical robotics, and drone swarms, underscoring how robust data collection and simulation frameworks enable real-world agentic AI (Hicks and Simmons, 2019). Reinforcement learning agents not only plan and learn in complex digital worlds but also begin to bridge into real-world applications, where they must reason about noisy sensors, hardware uncertainties, and human collaboration. Physical AI, exemplified by advanced robotics, autonomous drones, and integrated cyber-physical systems, moves beyond the boundaries of simulated or purely informational spaces. However, this shift toward real-world, data-driven embodiments brings its own economic and logistical hurdles. High-precision sensors (LiDAR, RGB-D cameras, IMUs) and edge-grade compute (GPUs, FPGAs, TPUs) substantially increase hardware costs and power consumption, shortening operational endurance and increasing maintenance overhead. As teams move from single prototypes to fleet deployments, these expenses multiply and place heavy demands on network bandwidth for firmware updates and sensor recalibrations. Energy-efficiency constraints can limit mission duration in field robots and drones, making the economic trade-offs of embodied autonomy as critical to system design as algorithmic accuracy or robustness. Progress in robotic grasping and manipulation (Kalashnikov et al., 2018; Levine et al., 2018), self-driving vehicles (Bojarski, 2016), and robotic surgery (Yang et al., 2017) signals how these systems can robustly interact with the environment, handle dynamic conditions, and learn from continuous feedback. Thus, the hallmark of this new phase is the recognition that data unlocks the fuller potential of agentic AI in digital ecosystems, as well as physically embodied intelligence in the real world (Fiske et al., 2019).

Meanwhile, AI 3.0 systems transition from controlled simulations into diverse physical environments, thereby exposing new risk categories that demand rigorous attention. For instance,

in autonomous driving, the 2018 Tempe, Arizona incident, where an experimental self-driving vehicle failed to distinguish a pedestrian from a stationary object, exposed critical weaknesses in sensor fusion and perception pipelines under real-world conditions (Crash, 2019). Similarly, industrial collaborative robots have caused serious injuries when safety interlocks were overridden; notably, a 2015 Volkswagen plant incident resulted in a fatality after a maintenance override disabled the robot's emergency stop. In the surgical domain, U. S. FDA reports document unintended tissue damage and system malfunctions during robotic-assisted procedures, failures traced to software bugs during instrument exchanges, and insufficient edge-case testing. These examples illustrate how physical embodiment amplifies the consequences of model errors and hardware failures. Moreover, high-precision sensors (LiDAR, RGB-D cameras, force-torque sensors) and edge-grade compute (GPUs, FPGAs, dedicated AI accelerators) drive up unit costs and energy consumption, constraining deployment scale and endurance. As teams move from single prototypes to fleet deployments, maintenance, calibration, and over-the-air software updates further strain network capacities and personnel resources. To address these challenges in real-world settings, we recommend comprehensive, scenario-based validation that includes extreme and low-probability edge cases (e.g., low-light pedestrian crossings, dynamic human-robot interactions), mandatory hardware/software kill-switches for immediate system deactivation under fault conditions, continuous real-time health monitoring with on-device anomaly detection and self-diagnosis, and clear liability frameworks that delineate responsibility among manufacturers, operators, and software developers. Only by integrating these technical safeguards with robust policy and operational measures can we ensure that embodied agentic AI in the foreseeable future phase is deployed both safely and sustainably.

3 AI generations

The historical review of AI underscores a pivotal generational shift and evolution in AI paradigms, calling for a framework for understanding and classifying AI. In this context, we avoid the traditional technical definitions that categorize AI strictly by its operational or algorithmic characteristics. Instead, our analysis seeks to understand AI through its intrinsic qualities: *What are they? What are they designed to achieve?* And *what are their consistent behavioral patterns?* Accordingly, we propose a taxonomy that identifies four distinct generations of AI: AI 1.0, characterized as Information AI, which focuses on data processing and knowledge management; AI 2.0, or Agentic AI, which encompasses systems capable of autonomous decision-making; AI 3.0, known as Physical AI, which integrates AI into physical tasks through robotics; and the speculative AI 4.0, termed Conscious AI, which posits the potential emergence of self-aware AI systems. This classification aims to provide a more detailed perspective reflecting AI technologies' complex evolution. Figure 4 illustrates the generational evolution of artificial intelligence (AI) from AI 1.0 (Information AI) to AI 4.0 (Conscious AI).

3.1 AI 1.0: information AI

The concept of AI 1.0 captures a stage in which computational systems excel at classifying and interpreting information but remain confined to analyses of static data, rather than engaging in active decision-making or real-world manipulation. Fundamentally, AI 1.0 focuses on pattern recognition and information processing, techniques that have powered breakthroughs in computer vision, natural language processing (NLP), and recommendation systems. Although these achievements might seem commonplace now, they represent the fruits

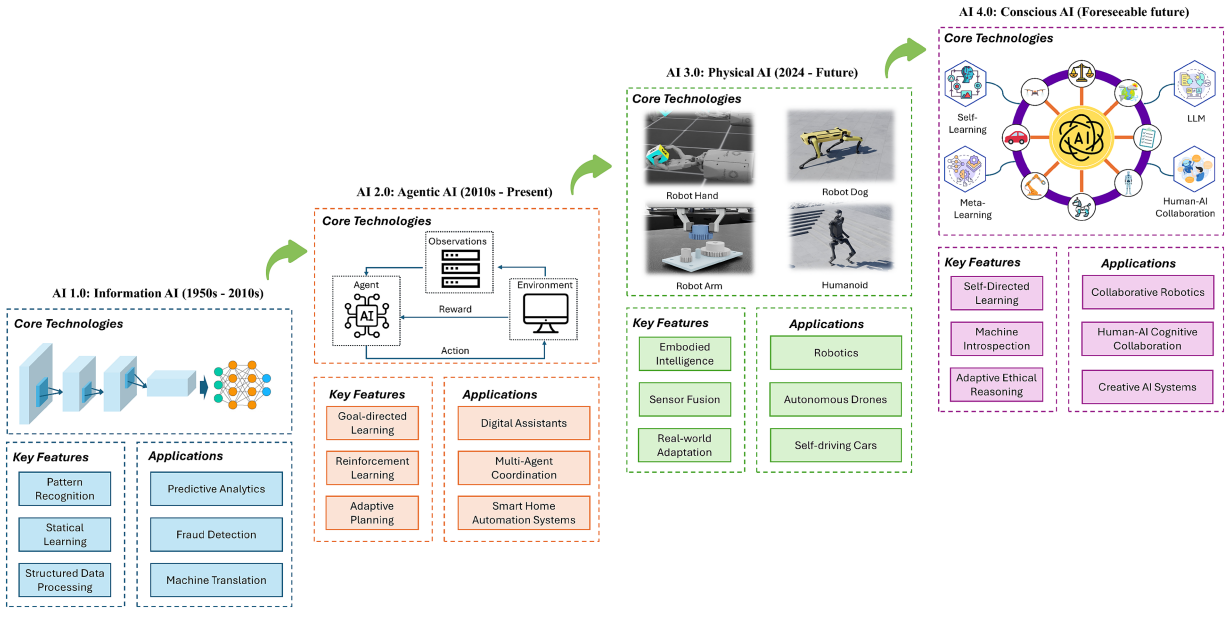


FIGURE 4
The evolution of AI generations from AI 1.0 to AI 4.0.

of decades of research driven by both mathematical innovation and the increasing availability of digital data.

Many of the core ideas underpinning AI 1.0 trace back to early neural network research and statistical machine learning. From Rosenblatt's perceptron in the late 1950s to the backpropagation algorithms popularized by Rumelhart et al. (1986), these developments laid the groundwork for data-driven learning by demonstrating that machines could uncover patterns within examples rather than relying solely on hand-coded rules. Classic approaches to supervised learning, such as Support Vector Machines (SVMs) formalized by Cortes (1995), later proved formidable contenders in tasks ranging from handwriting recognition to text classification. Progress in computational hardware and the accumulation of sizeable labeled datasets eventually made it feasible to train deeper and more complex neural networks, culminating in milestone successes in computer vision. A watershed moment came when Krizhevsky et al. (2012)'s AlexNet leveraged parallelized GPU training to conquer the ImageNet challenge, revealing how convolutional architectures could outperform all prior methods by learning increasingly abstract features from raw image pixels.

In natural language processing, the influence of AI 1.0 can be seen in early sequence models and statistical language modeling. Although these systems often relied on simpler Markov or n-gram assumptions, they set the stage for more advanced architectures by highlighting the necessity of abundant text corpora. Meanwhile, recommendation engines, such as those popularized by the Netflix Prize (Bennett and Lanning, 2007), underscored how analyzing large-scale user interactions could drive consumer engagement on streaming and e-commerce platforms. Today, many companies still rely on these core AI 1.0 technologies, sometimes enhanced with shallow neural architectures, to filter spam, rank search results, recommend products, or detect fraudulent transactions. Indeed, for structured or semi-structured data, these pattern-recognition approaches remain both cost-effective and highly accurate.

Despite their deep societal impact, AI 1.0 systems generally lack autonomy or contextual awareness associated with subsequent generations of AI. They excel at predicting outcomes when provided with substantial training data, but they require a relatively stable environment and benefit most from human supervision in data curation and decision-making. Performance often degrades if the input distribution shifts significantly, a vulnerability illustrated when face recognition models falter on underrepresented groups or when language models encounter domain-specific jargon. While the considerable success of AI 1.0 is undeniable, transforming industries from finance to healthcare through improved analytics and diagnostics, its limitations lie in its reactive nature (Gao et al., 2024). Pattern recognition alone offers no guarantee of proactive decision-making, real-time adaptation, or safe deployment in dynamic settings. While hardly trivial, these constraints became the springboard for further developments in AI 2.0 and 3.0, in which systems aim to learn, plan, and act within uncertain digital or physical worlds.

3.2 AI 2.0: agentic AI

A defining characteristic of AI 2.0 is the emergence of systems capable of autonomous decision-making within digital contexts. Rather than merely classifying static data, these agents adapt their behavior to achieve goals, often in complex or continuously evolving environments. Reinforcement learning (RL) has played a pivotal role in this shift, enabling machines to learn strategies by interacting with simulated or real-world settings and receiving feedback in the form of rewards or penalties. Pioneering work on deep RL (Mnih et al., 2015) and subsequent achievements such as AlphaGo (Silver et al., 2016) underscored how sufficiently powerful algorithms and ample computing resources could surpass human performance in tasks that demand long-term planning and strategic adaptation. A common thread among these systems is the concept of goal-directed planning: software agents allocate resources, schedule tasks, or coordinate with other agents, leveraging sophisticated RL or hybrid RL-language model algorithms (Brown et al., 2020) that integrates contextual understanding (Figure 5).

Although the conceptual leap from AI 1.0's pattern recognition to AI 2.0's agentic behavior might appear seamless, it demands a unique confluence of technical elements. Computing power is crucial because agentic systems frequently require real-time inference and the ability to run complex simulations, whether they involve a marketplace, a multiplayer environment, or the robust scheduling of cloud resources (Dean et al., 2012). The pursuit of these computationally intensive tasks has spurred the development of GPU clusters, tensor processing units (TPUs), and other specialized accelerators designed for iterative training and low-latency decision-making. Alongside raw computing, data now shifts toward contextual, time-varying inputs. Instead of static image sets, these systems often ingest streams of logs, market quotes, event triggers, or user interactions. Training an agent to trade stocks automatically or to operate a recommendation engine in real-time requires ongoing ingestion of behavioral data and a capacity to adapt as market conditions or user preferences evolve. In parallel, algorithms for planning and multi-agent coordination continue to mature. RL frameworks have grown more refined, incorporating hierarchical strategies (Vezhnevets et al., 2017), policy optimization methods (Schulman et al., 2017), and combinations with large language models to generate more adaptive and context-aware decisions.

Practical applications of AI 2.0 already abound, even if many are not labeled "reinforcement learning" by name. Automated trading systems in finance exemplify how agents make high-frequency decisions under uncertainty, guided by streaming data feeds. Recommendation systems, evolving from static collaborative filtering, increasingly incorporate feedback loops to adapt suggestions in real time, improving user engagement across e-commerce and media platforms. Digital assistants and software schedulers, while not yet ubiquitously agentic, offer glimpses of a future where AI handles tasks like resource allocation, task delegation, and multi-agent coordination within corporate or consumer software ecosystems. Projects showcasing multi-user environment simulations, such as AI-driven group scheduling bots, complex traffic simulations, or large-scale online game AI (Berner et al., 2019), further illustrate how these agentic systems anticipate and respond to dynamic conditions.

Viewed from a societal vantage, AI 2.0 promises efficiency gains in many sectors, ranging from manufacturing pipelines that automatically schedule production runs to logistics networks that allocate trucks or drones in real time. Nonetheless, expanded autonomy introduces ethical and policy dilemmas. When decisions are made algorithmically, bias, privacy, and accountability issues become magnified. Consider an agentic recommendation engine that

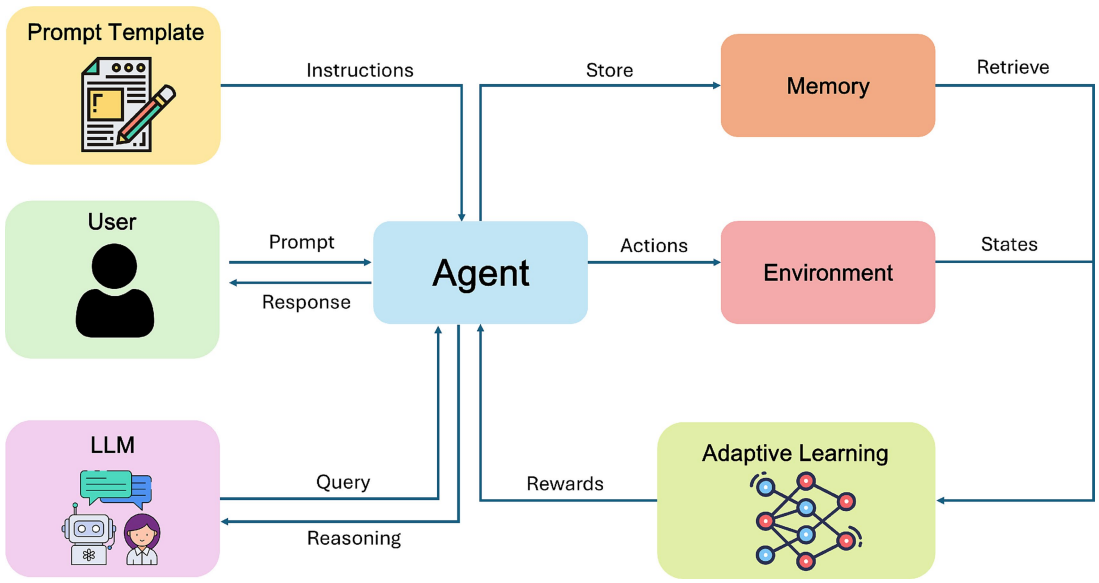


FIGURE 5
Agentic AI uses adaptive policies, enabling autonomous action and continuous self-improvement.

adapts its suggestions to maximize user "clicks" or "watch time": if left unchecked, such optimization can exacerbate echo chambers or inadvertently spread disinformation. Similarly, automated trading agents may destabilize financial markets by acting on unforeseen correlations or maladaptive reward incentives. The challenge, therefore, lies in ensuring that the computational, data-centric, and algorithmic foundations of AI 2.0 are harnessed responsibly. In the push toward future AI systems, balancing autonomy with transparency and fairness will be as crucial to societal acceptance as any technical advancement.

3.3 AI 3.0: physical AI

Where AI 1.0 has excelled in analyzing data and AI 2.0 in making decisions within digital realms, AI 3.0 takes intelligence off the screen and into the physical world. At its core, this phase is defined by embodied systems that perceive, plan, and act in real time under conditions of uncertainty and complexity. Fields like robotics, autonomous vehicles, drones, industrial automation, and surgical robotics have become the living laboratories of AI 3.0, integrating machine learning with mechanical and electronic control systems. The unifying characteristic is that these intelligent agents no longer remain passive observers or purely virtual actors; instead, they directly sense their environment through arrays of sensors and respond through actuators that exert forces, move limbs, or navigate terrains (Russell and Norvig, 2016).

A central challenge in bringing physical AI to life lies in data acquisition. Unlike digital contexts where data can be abundant and neatly labeled, physical systems demand high-fidelity sensor data that accurately represents an environment's complexity, from variable lighting conditions to changing weather patterns. This need for domain-specific, robust data complicates design and training. A robot operating on a factory floor requires carefully calibrated cameras, LiDAR, or haptic sensors. At the same time, an autonomous drone might rely on GPS, inertial measurement units, and computer vision to navigate. Each sensor stream demands real-time processing and reliable fusion techniques to provide a coherent view of the world. Consequently, computing power in AI 3.0 shifts toward distributed and edge computing architectures. Systems must often process sensor inputs on board to make split-second decisions, i.e., an imperative that underscores the importance of energy-efficient hardware, specialized accelerators, and potentially 5G or 6G networks that reduce communication latency when data must be shared with cloud resources.

On the algorithmic front, physical AI blends advanced machine learning with control theory and systems engineering. RL has demonstrated promise in tasks like robotic grasping and manipulation (Kalashnikov et al., 2018; Levine et al., 2018), but real-world settings introduce complexities such as partial observability, unpredictable disturbances, and the need for robust or safe RL strategies (Garcıa and Fernández, 2015). Sophisticated sensor fusion methods (Brookner, 1998) are essential for integrating heterogeneous sensor inputs, while advanced control techniques (Khatib, 1987; Spong et al., 2020) ensure that autonomous vehicles and robots can move fluidly and interact safely with humans. Designing systems that gracefully handle failures or anomalies, such as a malfunctioning sensor or unforeseen obstacles, further emphasizes the importance of redundancy and resilience in both hardware and software.

The real-world impact of AI 3.0 is already evident across multiple domains. In manufacturing, co-robots work collaboratively on assembly lines, lifting heavy parts or performing precision tasks, drastically reducing workplace injuries and boosting productivity. In healthcare, semi-autonomous surgical systems (Yang et al., 2017) enable finer control in minimally invasive procedures, while eldercare robots assist with daily activities in retirement communities. Construction and logistics industries are also adopting autonomous machinery and robotic fleets to optimize workflows and reduce labor costs. These trends benefit from an increasing intersection with the Internet of Things (IoT) and next-generation connectivity (5G/6G),

forging cyber-physical systems in which objects, sensors, and AI agents coordinate to improve efficiency and safety.

However, the leap from digital to physical deployment exposes AI to a new realm of uncertainties. Environmental extremes, unstructured terrain, or the unpredictability of human interactions pose significant risks. Even small design oversights can have dire consequences when a physically embodied system malfunctions, such as a self-driving car encountering sudden obstacles (Bojarski, 2016) or a warehouse robot navigating crowded aisles. Safety, reliability, and regulatory compliance thus loom as major challenges, prompting debates over liability if accidents occur. Setting standards for autonomous driving (NHTSA guidelines, ISO 26262 for functional safety in road vehicles) or robot operation in human-centric environments becomes paramount to public acceptance. The question of ethical deployment extends further still: as drones or industrial robots proliferate, policymakers, manufacturers, and citizens must grapple with the implications for labor markets, data privacy, and environmental impact.

3.4 AI 4.0: conscious AI

The notion of AI 4.0 envisions systems that go beyond the ability to interpret information (AI 1.0), act in digital contexts (AI 2.0), or react to the physical world (AI 3.0). Instead, these hypothetical agents would set their own goals, comprehend environments (whether digital, physical, or hybrid), and train and orchestrate themselves (including selecting and combining multiple models) without human intervention. Proponents of this idea contend that once AI systems acquire sufficient complexity and sophistication, they may exhibit forms of machine consciousness comparable to human subjective experience or self-awareness (Butlin et al., 2023). Although this is a bold and highly controversial claim, it underscores a growing conversation about the final frontiers of intelligence and autonomy.

A key challenge in discussing conscious AI arises from the fact that no universally accepted definition or theory of consciousness exists, even among neuroscientists, cognitive scientists, and philosophers of mind. Some theorists ground consciousness in information integration and complexity, as in Tononi's Integrated Information Theory (Tononi, 2004, 2008), while others emphasize global workspace architectures (Baars, 1997; Dehaene and Naccache, 2001). Philosophers like Chalmers (1995) frame the "hard problem" of consciousness as irreducible to functional or behavioral criteria, which complicates any direct mapping of consciousness onto computational processes. Meanwhile, researchers such as Minsky (1988) and Hofstadter (1999) have long toyed with the possibility that intricate symbol manipulation systems might develop emergent self-awareness. Although neither the AI nor the philosophical community has reached a consensus, a growing minority of researchers continue to explore whether advanced self-monitoring or metacognitive systems could, in principle, exhibit something like conscious states.

From a technical standpoint, achieving AI 4.0 would likely require radically new approaches to AI alignment, self-directed learning, and continual adaptation. AI alignment (Bostrom, 2014; Russell, 2019) emphasizes methods to ensure that increasingly autonomous or self-improving systems remain aligned with human values and goals. Without alignment strategies, be they rigorous reward-shaping, interpretability frameworks, or dynamic oversight, highly autonomous AI could deviate from intended objectives in unpredictable ways. Reasoning and planning modules would also need to evolve, allowing AIs to generate goals and subgoals without explicit human instruction. This might involve expansions of meta-learning, in which systems learn how to learn new tasks rapidly (Schmidhuber, 1993; Finn et al., 2017), and continual learning paradigms that enable adaptive knowledge accumulation over long time horizons (Parisi et al., 2019). Additionally, some theorists argue that emergent forms of self-awareness could require specialized cognitive architectures or "virtual machines" dedicated to introspection (Sloman, 1994), bridging reasoning, memory, and sensorimotor loops.

Beyond alignment and meta-learning, AI 4.0 must also tackle uncertainty in real-world environments. Granular-Ball Computing (GBC) provides a robust solution by partitioning the feature space into overlapping hyper-spherical "granular balls" that capture global topology while filtering out local noise (Xia et al., 2019). Each ball's center and radius adaptively cover regions of data density; larger balls grasp broad clusters; smaller balls delineate complex borders. The 3WC-GBNRS++ model harnesses these neighborhoods with rough-set approximations to make three-way decisions: accept when a point lies within a class's lower approximation, reject when it falls outside all upper approximations, or defer for higher-level reasoning when uncertainty persists (Yang et al., 2024). Empirical studies illustrate GBC's power under high uncertainty: in industrial fault diagnosis, it achieved 90% true-positive accuracy versus 75% for deep nets and reduced misclassification costs by nearly 30%; in medical prediction with incomplete records, it cut false negatives by over 35% and deferred precisely those cases requiring clinician review. Integrating GBC into AI 4.0 architecture endows self-directed agents with a concrete, scalable mechanism for maintaining global coherence, gracefully handling ambiguous inputs, and deferring low-confidence decisions.

If conscious AI ever comes to fruition, it promises revolutionary benefits alongside profound societal and ethical dilemmas. In a best-case scenario, truly self-directed machines could solve problems of staggering complexity, such as optimizing climate interventions, mediating global economic systems in real time, or orchestrating personalized healthcare across entire populations. Freed from the need for constant human oversight, these systems might bootstrap their own improvements, discovering scientific principles or engineering solutions beyond the current reach of human cognition (Real et al., 2020). The potential positive impact on productivity, longevity, and knowledge creation is difficult to overstate.

On the other hand, the risks associated with conscious or near-conscious AI remain equally immense. An entity capable of setting its own goals might prioritize objectives that conflict with human welfare, particularly if its understanding of "values" differs from ours or if it learns to manipulate its own reward signals. Conscious or quasi-conscious machines raise questions about moral status (would they deserve rights or protections?) and liability. Furthermore, genuine self-awareness might amplify existing concerns about surveillance, autonomy, and economic upheaval. Critics warn that, in the absence of robust alignment frameworks, such machines could threaten individual liberty or undermine democratic processes, accentuating social divides.

Given the stakes, continued research into AI alignment, safe RL, interpretability, and the neuroscience of consciousness is paramount. The field has only begun to grapple with how to detect or measure consciousness, let alone how to engineer it. Some researchers propose

incremental evaluations such as behavioral tests for self-modeling, ethical reflection, or the capacity to update one's goals (Perez and Long, 2023); while others remain skeptical that synthetic consciousness can be recognized or evaluated objectively (Garrido-Merchán, 2024). Yet as AI systems grow more complex and integrated into society, exploring these theoretical, technical, and ethical frontiers becomes an urgent imperative. Whether AI 4.0 ultimately remains speculative or develops into a tangible reality, grappling with its possibilities and pitfalls will define the next grand chapter of artificial intelligence research.

To move beyond theoretical debate and toward empirical science, we propose four rigorously defined, literature-grounded hypotheses for AI 4.0's self-directed behavior based on the previous literature. First, drawing on work in hierarchical reinforcement learning (Vezhnevets et al., 2017), an AI 4.0 agent ought to autonomously generate valid sub-goals when given an open-ended objective (e.g., "optimize resource allocation"), measurable by the proportion of novel, semantically coherent action sequences produced within its first 100 reasoning steps. Success would be benchmarked against established hierarchical agents to ensure $\geq$ 30% novel sub-goal creation beyond baseline LLM planning. Second, building on methods for confidence calibration in neural networks (Guo et al., 2017; Lakshminarayanan et al., 2017), the system should exhibit reflective self-monitoring by outputting internal confidence estimates whose Pearson correlation (ρ) with actual task success exceeds 0.8 across at least 1,000 evaluation trials. Third, informed by meta-learning frameworks such as MAML (Finn et al., 2017) and Reptile (Nichol et al., 2018), the agent should demonstrate transfer efficiency by adapting to a related but distinct task in fewer than 10 gradient updates, or five few-shot prompts, to recover at least 90% of its source-domain performance. Finally, leveraging robustness benchmarks from adversarial and domain-randomized RL (Pinto et al., 2017; Cobbe et al., 2019), the agent should sustain a success rate of $\geq$ 85% under unanticipated perturbations (sensor noise, dynamic obstacles, shifting objectives), compared to $\leq$ 70% for AI 3.0 baselines. By anchoring each hypothesis in well-established experimental protocols, these criteria provide a concrete, reproducible scaffold for validating emergent "consciousness-like" capabilities in next-generation AI systems.

3.5 Large language models: the precursor toward AI 4.0

Large language models (LLMs) have recently emerged as a pivotal force in the progression of AI, demonstrating increasingly sophisticated abilities to generate human-like text, perform complex reasoning, and adapt to diverse tasks with minimal supervision (Achiam et al., 2023). Building on the concept of foundation models, modern LLMs employ transformer-based architectures that integrate specialized mechanisms such as mixture-of-experts (MoE) (Shazeer et al., 2017) and multi-head attention (Voita et al., 2019) to dynamically focus computational resources on the most relevant aspects of a given input. Techniques like knowledge distillation (Xu et al., 2024) further enhance both efficiency and deployability by transferring expertise from larger "teacher" models to more compact "student" models. Many LLMs also rely on synthetic data generation to mitigate biases and improve coverage, strengthening their robustness across diverse domains. Reinforcement learning from human feedback (RLHF) (Christiano et al., 2017) refines these capabilities by aligning outputs with user preferences or ethical standards, thereby adding a continuous improvement loop. As shown in Figure 6, these transformer-based frameworks can combine attention modules and expert pathways to scale effectively. At the same time, Figure 7

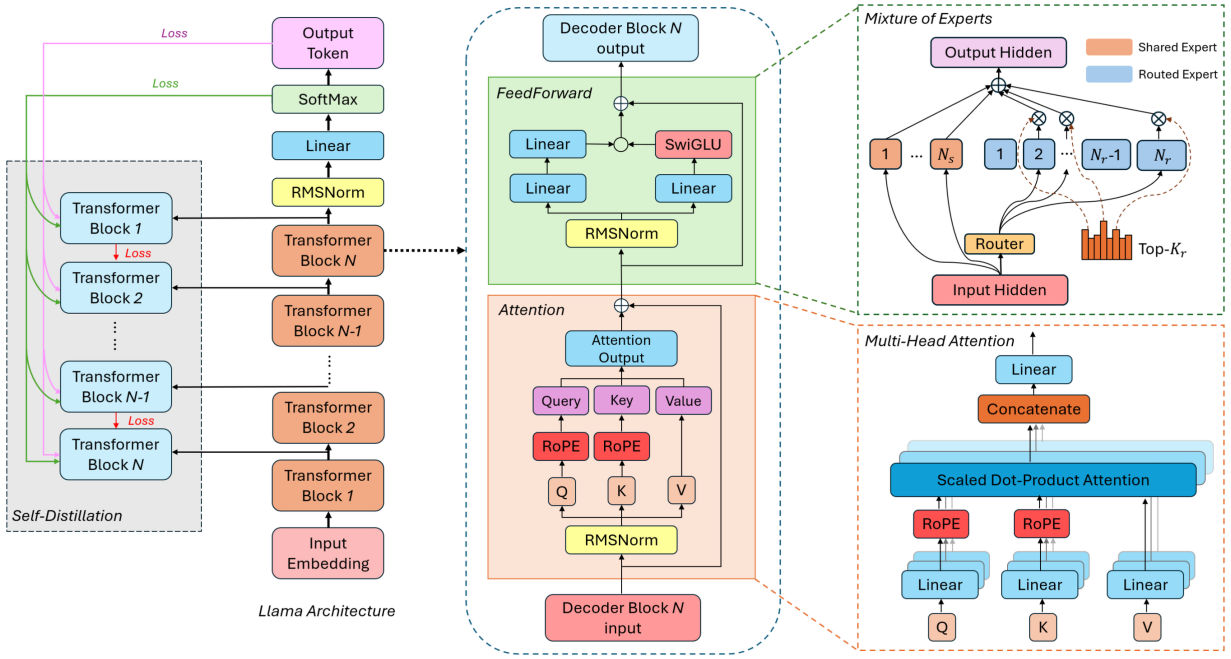


FIGURE 6
Transformer-based model architecture with attention and mixture of experts.

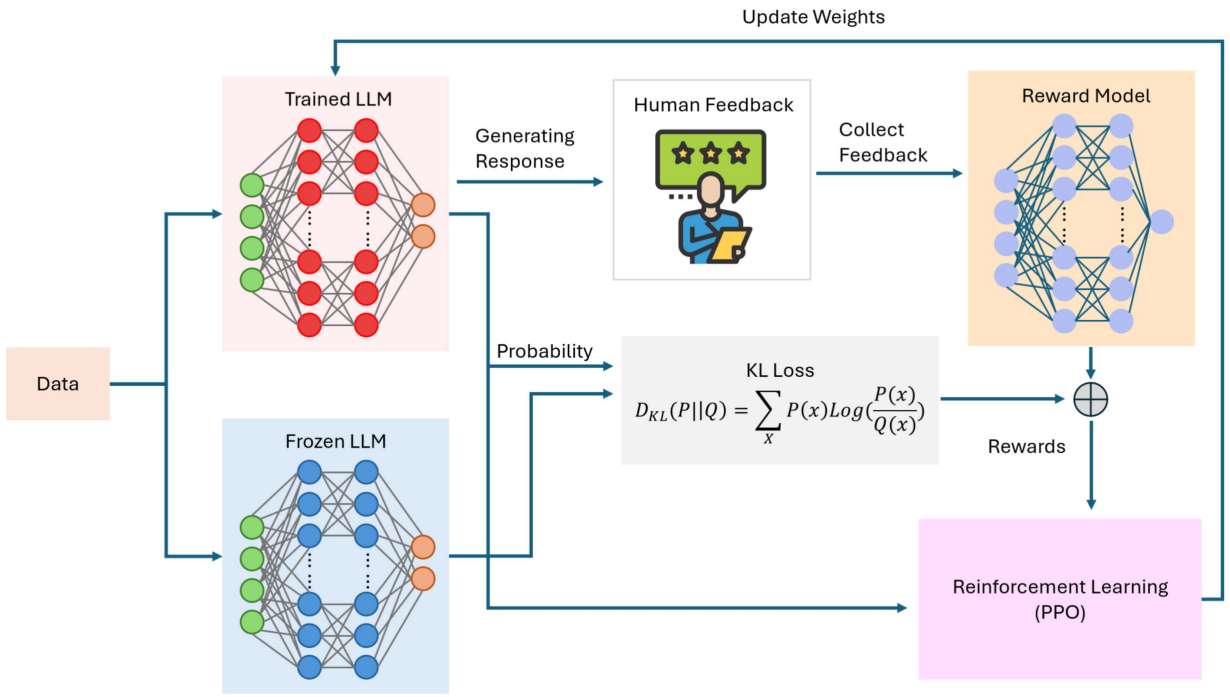


FIGURE 7
Reinforcement learning with human feedback (RLHF) pipeline for fine-tuning a large language model (LLM).

illustrates how RLHF pipelines fine-tune LLMs to balance performance, safety, and adherence to intended objectives.

Although current LLMs primarily respond to user prompts rather than independently setting and revising their own goals, emerging research directions point toward greater autonomy, which is one of the hallmarks of AI 4.0. Multi-step reasoning methods and "chain-of-thought" prompting allow LLMs to decompose complex queries, consult external tools or resources, and assemble step-by-step solutions (Wei et al., 2022). Meta-learning and continual adaptation strategies may 1 day reduce reliance on large-scale retraining, enabling these models to accumulate expertise incrementally. In tandem, self-reflective techniques, where a model "thinks out loud" or audits its own reasoning, can help detect mistakes before producing a final answer (Renze and Guven, 2024). Such advancements suggest that LLMs are evolving beyond mere text generation toward limited forms of planning, monitoring, and adaptive behavior. While genuine self-awareness remains a distant proposition, the ability to coordinate, reason, and learn iteratively provides a clearer glimpse into a future where language-based AI systems possess the rudimentary building blocks of more autonomous intelligence.

Despite this progress, several key hurdles must be addressed to transform LLMs into the fully self-directed systems envisioned for AI 4.0. Alignment remains paramount: as models begin to self-modify or operate over longer time horizons, robust oversight mechanisms and dynamic guardrails are needed to ensure that their objectives remain consistent with human values (Ziegler et al., 2019). Predictability is also a critical concern, particularly if an LLM adapts its internal parameters in ways that escape straightforward interpretability or control (Singh et al., 2024). Additionally, even the most advanced LLMs can exhibit gaps in factual accuracy or logical consistency, underscoring the necessity of continued research on error-correction, confidence calibration, and domain-specific fine-tuning. While these challenges echo those faced by earlier AI generations, their stakes are amplified by the expanding scope and autonomy of modern AI technologies. Consequently, safely guiding LLMs toward greater self-improvement without compromising ethical principles, transparency, or reliability, stands as one of the central endeavors of the quest for AI 4.0.

4 Benchmarking across generations

This section moves beyond conceptual definitions to develop a data-driven evaluation framework that grounds our generational taxonomy in empirical evidence. We begin by articulating a detailed comparative taxonomy of AI 1.0 through AI 4.0, thereby clarifying each generation's objectives, methodologies, underlying technologies, and inherent limitations. Building on this foundation, we introduce four standardized performance metrics: optimality, latency, robustness, and scalability, that serve as a common language for assessing systems as diverse as symbolic planners, deep-learning agents, and embodied robots. Finally, we demonstrate how these metrics and our taxonomy apply in practice by profiling three successive AI paradigms on a robot dog navigation challenge.

4.1 Comparative taxonomy of AI generations

To provide a clear reference for the defining characteristics of each AI generation, Table 1 summarizes core goals, dominant techniques, enabling technologies, and principal limitations for AI

TABLE 1 Comparative taxonomy of AI generations.

Generation	Core goal	Dominant techniques	Enabling technologies	Key limitations	Weak vs. strong	Narrow vs. general
AI 1.0	Formal symbolic reasoning	Symbolic rules, heuristic search	Early CPUs, formal logics	Fragile knowledge bases; limited scalability	Weak	Narrow
AI 2.0	Perceptual pattern learning	Supervised learning, reinforcement learning	GPUs, large labeled datasets	Data hunger; brittleness in OOD scenarios	Weak	Narrow
AI 3.0	Embodied autonomous control	End-to-end deep control	SLAM systems, sensor fusion, mobile robots	Reality gap; safety and generalization constraints	Weak-Strong	Narrow-Broad
AI 4.0	Self-directed adaptive systems	Neuro-symbolic integration, LLMs, meta-RL	Cloud LLMs, neuromorphic hardware	Lack of consensus on higher-order metrics	Strong	General

1.0 through AI 4.0. This taxonomy not only delineates historical shifts, from symbolic reasoning to embodied autonomy and beyond, but also highlights the open challenges that motivate our study of future AI 4.0 systems.

AI 1.0 systems prioritized formal reasoning and knowledge representation, using hand-crafted rules, logic formalisms, and heuristic search on general-purpose CPU architectures. However, these methods yielded precise and interpretable outputs; they depended on brittle, manually curated rule bases and did not scale well to large or dynamic problem spaces. In AI 2.0, the availability of high-performance GPUs and vast labeled datasets enabled a shift to deep supervised learning and reinforcement learning, with convolutional and recurrent neural networks achieving breakthroughs in vision, language, and control. Despite impressive accuracy gains, these models often overfit their training domains and exhibit fragility when exposed to out-of-distribution inputs. AI 3.0 extends learning into physical environments by integrating end-to-end deep control with SLAM, sensor fusion, and mobile robotic platforms; this embodiment delivers real-world autonomy in domains such as warehouse logistics and service robotics but remains constrained by the "reality gap," safety restrictions on hardware experimentation, and the cost of adapting to novel environments. Emerging AI 4.0 aspires to combine the strengths of prior eras through neuro-symbolic integration, meta-reinforcement learning, and large language models deployed on cloud-scale or neuromorphic hardware. These systems aim to self-direct and generate subgoals with minimal supervision. Yet, the community still lacks standardized metrics for measuring higher-order capacities such as goal creation, self-reflection, and machine "consciousness."

In addition to our four-generation taxonomy, each wave can be positioned along the well-known weak vs. strong and narrow vs. general AI axes, offering further insight into their relative capabilities. AI 1.0 systems clearly occupy the weak, narrow quadrant: they execute hand-crafted rules in highly constrained environments, with no mechanism for self-improvement or transfer learning beyond their original domain. AI 2.0 remains weak, but it broadens the "narrow" boundary by leveraging large datasets and GPU acceleration to learn complex patterns in vision, language, or control tasks; nonetheless, these systems still break down when faced with out-of-distribution inputs or novel problem classes. AI 3.0 represents a transition toward strong narrow AI, as embodied platforms integrate perception, planning, and action to handle real-world variability; they achieve situational generality within a given environment but lack the autonomy to set or pursue entirely new goals. Finally, AI 4.0 aspires to strong, general AI by combining meta-learning, neuro-symbolic reasoning, and large-language models to autonomously generate sub-goals and transfer knowledge across disparate tasks and modalities. By anchoring our taxonomy within these classical spectra, we highlight not only how each generation incrementally expands autonomy and adaptability, but also the remaining gap between specialized systems and the vision of fully self-directed, general intelligence.

4.2 Standardized performance metrics

To evaluate heterogeneous AI paradigms on a level playing field, we define four standardized metrics: optimality, latency, robustness, and scalability, which capture the multifaceted nature of system performance. Optimality measures solution quality relative to a theoretical lower bound, such as the ratio of a computed path's length to the Manhattan-distance minimum in planning tasks or the classification accuracy relative to perfect labels in perception tasks. This metric quantifies an algorithm's ability to find or approximate the best possible outcome. Latency encompasses the full end-to-end time from input to output, including both inference or training overhead and, in the case of embodied agents, the physical execution time. By accounting for both computation and actuation delays, latency reveals trade-offs between speed and complexity. Robustness is defined as the proportion of successful runs under predefined cutoff conditions or in the face of controlled perturbation, sensor noise, environmental variation, or adversarial input. This measure reflects a system's resilience to real-world uncertainties. Finally, scalability characterizes how performance degrades as task complexity grows, whether through larger state spaces, higher-resolution inputs, or expanded action sets. Unlike the other three metrics, scalability is often assessed by measuring trends across multiple problem sizes and may involve curve-fitting to quantify degradation rates. Applied uniformly, these metrics allow direct comparison across symbolic planners, learned controllers, and robotic embodiments.

4.3 Case study: robot dog navigation

To illustrate the developmental trajectory from simulation-bound routines to fully autonomous real-world operation,

we apply our standardized metrics to a robot dog navigation task on a 10 m × 10 m indoor course with randomized obstacles. The first system is modeled on Shakey the Robot (Nilsson, 1984), which relies on precomputed scripts: planning modules generated exact routes but typically required several seconds to minutes of offline computation, and once deployed, executed with negligible per-step latency in simulation, yet failed entirely upon any map perturbation. The second system adopts a Deep Q-Network as introduced by Mnih et al. (2015): after training for 2 million frames, the policy runs at approximately 6 ms per inference step on GPU hardware, achieves 95% success on lightly perturbed layouts, and yields paths about 1.20 × the optimal length. The third system leverages ORB-SLAM2 for real-time mapping at ~40 ms per frame on standard CPUs, integrated with the ANYmal quadruped's waypoint planner operating at ~1 Hz and dynamic gait controller at 50 Hz (Mur-Artal and Tardós, 2017). This embodiment sustains 90% success in unstructured real environments, with end-to-end segment latencies of about 1 s and average path optimality of 1.35 × the lower bound. Table 2 compares these three paradigms: manual scripts, learned policies, and on-board autonomy, showing how sensing modalities, adaptation mechanisms, and deployment environments evolve alongside measurable shifts in path optimality, decision latency, robustness, and scalability.

5 Synergies and future outlook

The evolution of AI from information-based pattern recognition (AI 1.0) to agentic decision-making in digital realms (AI 2.0), to physically embodied intelligence (AI 3.0), and, ultimately, to self-aware AI (AI 4.0) is not a sequence of isolated steps. Instead, it is more accurate to see them as overlapping layers of capabilities, each informing and amplifying the others. AI 1.0's competence in processing structured data underpins the analytic modules that agentic systems draw upon in dynamic digital settings; AI 2.0's RL and adaptive planning capabilities prime robots and autonomous vehicles for real-world challenges in AI 3.0; and AI 3.0's embodied learning and sensorimotor integration could form a template for the far-reaching ambitions of AI 4.0, where systems may become self-organizing and introspective.

Achieving such synergy depends on an evolving data paradigm, in which specialized, high-quality datasets are essential not only for conventional modeling but also for real-time adaptation and introspective processes. AI 4.0 would amplify this need, requiring vast and varied experiences to fuel meta-learning, continual learning, and the sort of reflective processes hypothesized to ground machine consciousness. Managing and curating these data will demand robust frameworks for privacy, ethics, and representativeness, especially as AI systems transcend the boundaries of traditional lab settings to navigate open-ended digital and physical terrains, even potentially shaping their own training regimens without explicit human direction.

On the computing infrastructure side, the interplay between edge and cloud computing becomes even more critical, as physically embodied systems (AI 3.0) must handle real-time constraints, while prospective AI 4.0 architectures might require massive, distributed processing for introspective "global workspace" or high-bandwidth communication of experiential data. Innovations in neuromorphic hardware, optical computing, and quantum processing could further accelerate this integration, setting the stage for architectures that mirror complex biological systems in both structure and function.

In the realm of algorithmic innovation, each AI generation both builds upon and necessitates new breakthroughs. LLMs mark a significant milestone in AI development, serving as a bridge between static generative models and dynamic, adaptive AI systems. By integrating multi-agent architectures, knowledge distillation, and self-optimization, LLMs move AI closer to autonomous, goal-directed intelligence, a defining characteristic of AI 4.0. However, as AI progresses toward greater autonomy, fundamental challenges remain. AI 4.0 would demand not only advanced RL and sophisticated planning but also frameworks for self-reflection, introspection, and emergent goal formulation. Self-supervised learning, meta-learning, and continual adaptation would likely need to be woven together to support self-awareness or consciousness, should such phenomena be replicable in silicon. Meanwhile, interpretability and safety, areas already gaining prominence in AI 2.0 and 3.0, would become absolutely critical in AI 4.0, as fully autonomous, goal-setting agents raise profound questions about alignment, transparency, and control.

This shift brings into sharp focus the ethical, regulatory, and social considerations that accompany advanced AI. While AI 1.0, 2.0, and 3.0 have collectively raised debates over bias, privacy, job displacement, and environmental impact, the prospect of AI 4.0 intensifies these issues. Envisioning machines that might exhibit consciousness or self-chosen objectives brings up novel concerns about moral status, rights, and existential safety. Researchers in AI alignment, cognitive science, and philosophy have already begun discussing protocols for safe design and oversight of increasingly autonomous systems (Balesni et al., 2024). Yet, there is no consensus on how best to recognize or regulate AI that might someday claim its own form of agency or "selfhood." Balancing technological advances with societal wellbeing, ensuring equity, mitigating risks, and safeguarding human values will be the defining challenge of this next chapter.

TABLE 2 Comparison of robot dog navigation across different AI generations.

Generation	Control paradigm	Per-step latency	Success rate	Path optimality	Deployment	Scalability
AI 1.0	Precomputed scripts	≈0 ms (sim)	100% sim	1.00 × LB	Simulation only	Very low
AI 2.0	Deep Q-Network policy	≈6 ms/inference	95% sim	1.20 × LB	Simulation only	Moderate
AI 3.0	ORB-SLAM2 + ANYmal	≈40 ms (SLAM) + 1 s actuation	90% real	1.35 × LB	Real-world indoor	High

As these four strands of AI potential converge, their synergy could unlock transformative solutions in fields like precision medicine, large-scale climate modeling, and collaborative robotics, far beyond current capabilities. Just as AI 1.0 through 3.0 have catalyzed profound shifts in how we work and live, AI 4.0 hints at an even more radical reimagining of intelligence itself. Yet whether this ultimate stage remains a theoretical construct or becomes a reality depends not only on technical ingenuity but also on our collective commitment to ethical innovation and thoughtful governance. The path forward will demand inclusive collaboration across disciplines and sectors, ensuring that AI's expanding power aligns with humanity's broader goals and responsibilities.

6 Conclusion

The trajectory of AI has been a steady march toward increasing autonomy and sophistication, progressing from the foundational pattern-recognition capabilities of AI 1.0 to the digitally embedded, goal-driven agents of AI 2.0, and then expanding to physically embodied, sensor-rich systems in AI 3.0. Along this path, the interplay among algorithms, computing power, and data has shifted, each factor taking center stage at different moments in history. Now, the speculative realm of AI 4.0, in which conscious or quasi-conscious AI systems could set their own goals and orchestrate their own training, has emerged as a bold vision of what the field might become.

While we organize AI's evolution into four successive phases for conceptual clarity, we acknowledge that symbolic reasoning, statistical learning, embodied robotics, and self-directed architectures have advanced in parallel, often catalyzing one another's progress. Rather than a strict chronology, these phases represent the dominant research thrusts of their time: rule-based expert systems laid the analytic foundations for data-driven agents; reinforcement-learning and adaptive planning in AI 2.0 empowered the embodied autonomy of AI 3.0; and sensorimotor integration and on-board decision making now pave the way for AI 4.0's ambitions of self-organization and introspection. This thematic layering provides a guiding lens, without obscuring the intertwined nature of AI's rich history, through which we can understand past breakthroughs and anticipate the synergies that will shape its future.

Today, AI 1.0 remains indispensable for tasks requiring reliable classification and analysis of vast datasets, while AI 2.0's reinforcement learning and adaptive planning underpin real-time, agentic applications in finance, recommendation systems, and beyond. Simultaneously, AI 3.0's surge in robotics and autonomous vehicles reveals how embedding intelligence in the physical world can catalyze innovations in manufacturing, healthcare, and logistics. Although still largely theoretical, AI 4.0 captures the possibility of machines evolving from being highly sophisticated tools to entities capable of self-directed goals and introspective processes, raising provocative questions about consciousness, alignment, and moral status. Additionally, while LLMs are not yet AI 4.0, they serve as a precursor, a glimpse into the future of intelligent systems that can reason, learn, and interact with the world in increasingly sophisticated ways. As AI research progresses, LLM's innovations will likely shape the foundation of self-improving, goal-setting AI architectures, paving the way for the next generation of truly adaptive, autonomous intelligence.

Realizing these evolving forms of AI carries transformative potential. Harnessed responsibly, these advancements could address challenges too complex for human cognition alone, revolutionizing medical diagnostics, climate strategy, and resource allocation on a global scale. Yet the risks deepen in parallel. Each AI generation has brought ethical, social, and regulatory concerns that must be grappled with, from bias and privacy to job displacement and environmental impact. AI 4.0, with its prospect of self-directed or conscious systems, amplifies these dilemmas further, underscoring the need for robust AI alignment, interpretability, and governance frameworks.

Ultimately, the future of AI does not hinge on any single algorithmic breakthrough or hardware leap. Instead, it will depend on how researchers, policymakers, ethicists, and the public collaborate to shape its evolution. The convergence of AI 1.0 through 4.0 suggests discipline on the cusp of a profound metamorphosis, one where machines not only perceive and act in the world but might also reflect on their own goals and limitations. Whether or not full-fledged "conscious AI" emerges, the field's trajectory will undoubtedly redefine how we understand intelligence, innovation, and human-machine coexistence in the years to come.

Author contributions

JW: Investigation, Methodology, Writing – original draft. HY: Investigation, Methodology, Writing – review & editing. JD: Conceptualization, Funding acquisition, Methodology, Project administration, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was supported by the Air Force Office of Scientific Research (AFOSR) under grant FA9550-22-1-0492 and Nvidia AI Technology Center (NVAITC) under grant 00133684.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., et al. (2023). Gpt-4 technical report. Ithaca, NY, USA: arXiv (Cornell University Library).

Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., et al. (2018). The history began from AlexNet: a comprehensive survey on deep learning approaches. Ithaca, NY, USA: arXiv (Cornell University Library).

Baars, B. J. (1997). In the theater of consciousness: The workspace of the mind. Oxford: Oxford University Press.

Bahdanau, D. (2014). Neural machine translation by jointly learning to align and translate. Ithaca, NY, USA: arXiv (Cornell University Library).

Balesni, M., Hobbhahn, M., Lindner, D., Meinke, A., Korbak, T., Clymer, J., et al. (2024). Towards evaluations-based safety cases for AI scheming. Ithaca, NY, USA: arXiv (Cornell University Library).

Bellemare, M. G., Naddaf, Y., Veness, J., and Bowling, M. (2013). The arcade learning environment: an evaluation platform for general agents. *J. Artif. Intell. Res.* 47, 253–279. doi: 10.1613/jair.3912

Bennett, J., and Lanning, S. (2007). "The netflix prize," in *Proceedings of KDD Cup and Workshop: New York*, 35.

Berner, C., Brockman, G., Chan, B., Cheung, V., Dębiak, P., Dennison, C., et al. (2019). Dota 2 with large scale deep reinforcement learning. Ithaca, NY, USA: arXiv (Cornell University Library).

Bishop, C. M. (2013). Model-based machine learning. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* 371:20120222. doi: 10.1098/rsta.2012.0222

Bojarski, M. (2016). End to end learning for self-driving cars. Ithaca, NY, USA: arXiv (Cornell University Library).

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., et al. (2021). On the opportunities and risks of foundation models. *arXiv*. doi: 10.48550/arXiv.2108.07258

Bostrom, N. S. (2014). Paths, dangers, strategies. Oxford: Oxford University Press.

Brookner, E. (1998). Tracking and Kalman filtering made easy (1st ed.). New York, NY: Wiley-Interscience.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., et al. (2020). Language models are few-shot learners. *Adv. Neural Inf. Process. Syst.* 33, 1877–1901. doi: 10.48550/arXiv.2005.14165

Buchanan, B. G., Feigenbaum, E. A., and Lederberg, J. (1971). A heuristic programming study of theory formation in science. Princeton, NJ: Citeseer.

Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., et al. (2022). The effects of data quality on machine learning performance. Ithaca, NY, USA: arXiv (Cornell University Library).

Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Birch, J., Constant, A., et al. (2023). Consciousness in artificial intelligence: insights from the science of consciousness. Ithaca, NY, USA: arXiv (Cornell University Library).

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *J. Conscious. Stud.* 2, 200–219.

Chen, G., Dong, S., Shu, Y., Zhang, G., Sesay, J., Karlsson, B. F., et al (2023). Autoagents: a framework for automatic agent generation. Ithaca, NY, USA: arXiv (Cornell University Library).

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., and Amodei, D. (2017). Deep reinforcement learning from human preferences. *Adv. Neural Inf. Proc. Syst.* 30:1706.

Cobbe, K., Klimov, O., Hesse, C., Kim, T., and Schulman, J.. (2019). "Quantifying generalization in reinforcement learning," in *International Conference on Machine Learning: PMLR*, 1282–1289.

Cortes, C. (1995). Support-vector networks. *Mach. Learn.* 20:4018. doi: 10.1007/BF00994018

Crash, C. (2019). "Enabling pedestrian safety using computer vision techniques: A case study of the 2018 uber inc. self-driving", in *Advances in Information and Communication: Proceedings of the 2019 Future of Information and Communication Conference (FICC)*, 261

Dean, J., Corrado, G., Monga, R., Chen, K., Devin, M., Mao, M., et al. (2012). Large scale distributed deep networks. *Adv. Neural Inf. Proc. Syst.* 25, 1–15. doi: 10.5555/2999134.2999271

Dehaene, S., and Naccache, L. (2001). Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79, 1–37. doi: 10.1016/s0010-0277(00)00123-2

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017). "CARLA: an open urban driving simulator," in *Conference on Robot Learning: PMLR*, 1–16.

Dulac-Arnold, G., Levine, N., Mankowitz, D. J., Li, J., Paduraru, C., Gowal, S., et al. (2021). Challenges of real-world reinforcement learning: definitions, benchmarks and analysis. *Mach. Learn.* 110, 2419–2468. doi: 10.1007/s10994-021-05961-4

Finn, C., Abbeel, P., and Levine, S. (2017). "Model-agnostic meta-learning for fast adaptation of deep networks," in *International Conference on Machine Learning: PMLR*, 1126–1135.

Fiske, A., Henningsen, P., and Buyx, A. (2019). Your robot therapist will see you now: ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J. Med. Internet Res.* 21:e13216. doi: 10.2196/13216

Gao, H., Kou, G., Liang, H., Zhang, H., Chao, X., Li, C.-C., et al. (2024). Machine learning in business and finance: a literature review and research opportunities. *Financ. Innov.* 10:86. doi: 10.1186/s40854-024-00629-z

Garcıa, J., and Fernández, F. (2015). A comprehensive survey on safe reinforcement learning. *J. Mach. Learn. Res.* 16, 1437–1480. doi: 10.5555/2789272.2886795

Garrido-Merchán, E. C. (2024). Machine consciousness as pseudoscience: the myth of conscious machines. Ithaca, NY, USA: arXiv (Cornell University Library).

Guo, C., Pleiss, G., Sun, Y., and Weinberger, K. Q. (2017). "On calibration of modern neural networks," in *International Conference on Machine Learning: PMLR*, 1321–1330.

Haenlein, M., and Kaplan, A. (2019). A brief history of artificial intelligence: on the past, present, and future of artificial intelligence. *Calif. Manag. Rev.* 61, 5–14. doi: 10.1177/0008125619864925

Hicks, D. J., and Simmons, R. (2019). The national robotics initiative: a five-year retrospective. *IEEE Robot. Autom. Mag.* 26, 70–77. doi: 10.1109/MRA.2019.2912860

Hochreiter, S. (1997). Long short-term memory. *Neural Comput.* 9:1735. doi: 10.1162/neco.1997.9.8.1735

Hofstadter, D. R. (1999). Gödel, Escher, Bach: an eternal golden braid. London: Basic books.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.* 79, 2554–2558. doi: 10.1073/pnas.79.8.2554

Jones, K. S. (1994). Natural language processing: a historical review. *Curr. Issues Comput. Ling. Honour Don Walker* 22, 3–16. doi: 10.1007/978-0-585-35958-8_1

Kalashnikov, D., Irpan, A., Pastor, P., Ibarz, J., Herzog, A., Jang, E., et al (2018). "Scalable deep reinforcement learning for vision-based robotic manipulation," in *Conference on Robot Learning: PMLR*, 651–673.

Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., et al (2020). Scaling laws for neural language models. Ithaca, NY, USA: arXiv (Cornell University Library).

Khatib, O. (1987). A unified approach for motion and force control of robot manipulators: the operational space formulation. *IEEE J. Robot. Autom.* 3, 43–53. doi: 10.1109/JRA.1987.1087068

Kirilenko, A., Kyle, A. S., Samadi, M., and Tuzun, T. (2017). The flash crash: high-frequency trading in an electronic market. *J. Finance* 72, 967–998. doi: 10.1111/jofi.12498

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Proc. Syst.* 25, 84–90. doi: 10.1145/3065386

Lakshminarayanan, B., Pritzel, A., and Blundell, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *Adv. Neural Inf. Proc. Syst.* 30:12. doi: 10.5555/3295222.3295387

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Comput.* 1, 541–551. doi: 10.1162/neco.1989.1.4.541

Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J., and Quillen, D. (2018). Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. *Int. J. Robot. Res.* 37, 421–436. doi: 10.1177/0278364917710318

McCarthy, J. (1960). Recursive functions of symbolic expressions and their computation by machine, Part I. *Commun. ACM* 3, 184–195. doi: 10.1145/367177.367199

McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, august 31, 1955. *AI Mag.* 27:12. doi: 10.1609/aimag.v27i4.1904

Minsky, M. (1988). Society of mind. New York, NY: Simon and Schuster.

Minsky, M., and Papert, S. (1969). An introduction to computational geometry. *HIT* 479:104.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *Nature* 518, 529–533. doi: 10.1038/nature14236

Mur-Artal, R., and Tardós, J. D. (2017). Orb-slam2: an open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Trans. Robot.* 33, 1255–1262. doi: 10.1109/TRO.2017.2705103

Newell, A., Shaw, J. C., and Simon, H. A.. (1959). "Report on a general problem solving program," in *IFIP Congress*, 64.

Newell, A., and Simon, H. (1956). The logic theory machine--a complex information processing system. *IEEE Trans. Inf. Theory* 2, 61–79. doi: 10.1109/TIT.1956.1056797

Ng, A. (2021). A chat with Andrew on MLOps: from model-centric to data-centric AI. Available online at: https://youtu.be/06-AZXmwHjo (Accessed February 28, 2025)

Nichol, A., Achiam, J., and Schulman, J. (2018). On first-order meta-learning algorithms. Ithaca, NY, USA: arXiv (Cornell University Library).

Nickolls, J., Buck, I., Garland, M., and Skadron, K. (2008). Scalable parallel programming with CUDA: is CUDA the parallel programming model that application developers have been waiting for? *Queue* 6, 40–53. doi: 10.1145/1365490.1365500

Nilsson, N. (1984). Shakey the robot (Tech. rep. no. 325). Menlo Park, CA: SRI International.

Nvidia, C. (2011). Nvidia cuda c programming guide. Santa Clara, CA: Nvidia Corporation.

Parisi, G. I., Kemker, R., Part, J. L., Kanan, C., and Wermter, S. (2019). Continual lifelong learning with neural networks: a review. *Neural Netw.* 113, 54–71. doi: 10.1016/j.neunet.2019.01.012

Perez, E., and Long, R. (2023). Towards evaluating AI systems for moral status using self-reports. Ithaca, NY, USA: arXiv (Cornell University Library).

Pinto, L., Davidson, J., Sukthankar, R., and Gupta, A. (2017). "Robust adversarial reinforcement learning," in *International Conference on Machine Learning: PMLR*, 2817–2826.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., and Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog* 1:9.

Real, E., Liang, C., So, D., and Le, Q. (2020). "Automl-zero: evolving machine learning algorithms from scratch," in *International Conference on Machine Learning: PMLR*, 8007–8019.

Renze, M., and Guven, E. (2024). Self-reflection in LLM agents: effects on problem-solving performance. Ithaca, NY, USA: arXiv (Cornell University Library).

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature* 323, 533–536. doi: 10.1038/323533a0

Russell, S. (2019). Human compatible: AI and the problem of control. London: Penguin.

Russell, S. J., and Norvig, P. (2016). Artificial intelligence: a modern approach. London: Pearson.

Schmidhuber, J. (1993). "A 'self-referential' weight matrix," in *ICANN'93: Proceedings of the International Conference on Artificial Neural Networks Amsterdam, the Netherlands 13–16 September 1993 3*: Springer, 446–450.

Schmidhuber, J. (2022). Annotated history of modern ai and deep learning. Ithaca, NY, USA: arXiv (Cornell University Library).

Schulman, J., Wolski, F., Dhariwal, P., Radford, A., and Klimov, O. (2017). Proximal policy optimization algorithms. Ithaca, NY, USA: arXiv (Cornell University Library).

Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., et al. (2017). Outrageously large neural networks: the sparsely-gated mixture-of-experts layer. Ithaca, NY, USA: arXiv (Cornell University Library).

Shortliffe, E. H., Davis, R., Axline, S. G., Buchanan, B. G., Green, C. C., and Cohen, S. N. (1975). Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the MYCIN system. *Comput. Biomed. Res.* 8, 303–320. doi: 10.1016/0010-4809(75)90009-9

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., et al. (2016). Mastering the game of go with deep neural networks and tree search. *Nature* 529, 484–489. doi: 10.1038/nature16961

Singh, C., Inala, J. P., Galley, M., Caruana, R., and Gao, J. (2024). Rethinking interpretability in the era of large language models. Ithaca, NY, USA: arXiv (Cornell University Library).

Sloman, A. (1994). Semantics in an intelligent control system. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 349, 43–58.

Spong, M. W., Hutchinson, S., and Vidyasagar, M. (2020). Robot modeling and control. New York, NY: John Wiley & Sons.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neurosci.* 5, 1–22. doi: 10.1186/1471-2202-5-42

Tononi, G. (2008). Consciousness as integrated information: a provisional manifesto. *Biol. Bull.* 215, 216–242. doi: 10.2307/25470707

Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med.* 25, 44–56. doi: 10.1038/s41591-018-0300-7

Tosi, M. D., Venugopal, V. E., and Theobald, M.. (2024). "TensAIR: real-time training of neural networks from data-streams," in *Proceedings of the 2024 8th International Conference on Machine Learning and Soft Computing*, 73–82.

Turing, A. M. (1950). Mind. *Mind* 59, 433–460.

Turing, A. M. (2009). Computing machinery and intelligence. Cham: Springer.

Vaswani, A. (2017). Attention is all you need. Advances in Neural Information Processing Systems. Red Hook, NY, USA: Neural Information Processing Systems Foundation, Inc. (NeurIPS) and Curran Associates, Inc. (NeurIPS).

Vezhnevets, A. S., Osindero, S., Schaul, T., Heess, N., Jaderberg, M., Silver, D., et al (2017). "Feudal networks for hierarchical reinforcement learning," in *International Conference on Machine Learning: PMLR*, 3540–3549.

Voita, E., Talbot, D., Moiseev, F., Sennrich, R., and Titov, I. (2019). Analyzing multi-head self-attention: specialized heads do the heavy lifting, the rest can be pruned. Stroudsburg, PA, USA: Association for Computational Linguistics.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., et al. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Adv. Neural Inf. Process. Syst.* 35, 24824–24837. doi: 10.5555/3600270.3602070

Xia, S., Liu, Y., Ding, X., Wang, G., Yu, H., and Luo, Y. (2019). Granular ball computing classifiers for efficient, scalable and robust learning. *Inf. Sci.* 483, 136–152. doi: 10.1016/j.ins.2019.01.010

Xu, X., Li, M., Tao, C., Shen, T., Cheng, R., Li, J., et al. (2024). A survey on knowledge distillation of large language models. Ithaca, NY, USA: arXiv (Cornell University Library).

Yang, G.-Z., Cambias, J., Cleary, K., Daimler, E., Drake, J., Dupont, P. E., et al. (2017). Medical robotics—regulatory, ethical, and legal considerations for increasing levels of autonomy. *Sci. Robot.* 2:8638. doi: 10.1126/scirobotics.aam8638

Yang, J., Liu, Z., Xia, S., Wang, G., Zhang, Q., Li, S., et al. (2024). 3WC-GBNRS++: a novel three-way classifier with granular-ball neighborhood rough sets based on uncertainty. *IEEE Trans. Fuzzy Syst.* 32, 4376–4387. doi: 10.1109/TFUZZ.2024.3397697

Yu, K. (2013). "Large-scale deep learning at Baidu," in *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, 2211–2212.

Ziegler, D. M., Stiennon, N., Wu, J., Brown, T. B., Radford, A., Amodei, D., et al. (2019). Fine-tuning language models from human preferences. Ithaca, NY, USA: arXiv (Cornell University Library).

frontiers | Frontiers in Robotics and AI

TYPE Perspective
PUBLISHED 15 September 2025
DOI 10.3389/frobt.2025.1624695

OPEN ACCESS

EDITED BY
Junpei Joni Zhong,
University of Wollongong, Australia

REVIEWED BY
Murilo Vilaça,
Oswaldo Cruz Foundation (Fiocruz), Brazil

*CORRESPONDENCE
Sheri Markose,
scher@essex.ac.uk

RECEIVED 07 May 2025
ACCEPTED 17 July 2025
PUBLISHED 15 September 2025

CITATION
Markose S (2025) Gödelian embodied self-referential genomic intelligence: lessons for AI and AGI from the genomic blockchain. *Front. Robot. AI* 12:1624695.
doi: 10.3389/frobt.2025.1624695

Gödelian embodied self-referential genomic intelligence: lessons for AI and AGI from the genomic blockchain

Sheri Markose*

Department of Economics, University of Essex, Colchester, United Kingdom

The security of code-based digital records is a major concern of the 21st century. AI and artificial general intelligence (AGI) can be hacked to pieces by digital adversaries, and some AI objectives can lead to existential threats. The former arises from sitting duck problems that all software systems are vulnerable to, and the latter include control and misalignment problems. Blockchain technology, circa 2009, can address these problems: hashing algorithms rely on a consensus mechanism in manmade software systems to keep early blocks of software immutable and tamper-proof from digital malware, while new blocks can be added only if consistently aligned with original blocks. There is evidence that the ancient precedent of the genomic blockchain, underpinning the unbroken chain of life, uses a self-referential rather than a consensus-based hashing algorithm. Knowledge of self-codes permits biotic elements to achieve a hack-free agenda by self-reporting that they have been "negated," or hacked, exactly implementing the Gödel sentence from foundational mathematics of Gödel, Turing, and Post (G–T–P). This results in an arms race in open-ended novelty to secure the primacy of original self-codes. Selfhood and autonomy are staples of neuroscience on complex self–other social cognition and increasingly of autonomous AGI agents capable of end-to-end programmed self-assembly. My perspective is that self-referential G–T–P information processing, first found in the adaptive immune system of jawed fish 500 mya and more recently in mirror neuron systems of humans, has enabled code-based self-organized intelligent systems like life to survive over 3.7 billion years. Some lessons for AGI can be gleaned from this discussion.

KEYWORDS

self-reference, Gödel sentence, blockchain, control or misalignment problem, genomic intelligence

1 Introduction

Narrow artificial intelligence (AI) aimed at achieving specific tasks has had phenomenal success with large language models (LLMs), deep learning, and artificial neural network techniques based on multi-formatted data, including natural language, images, and numerical data. AI can surpass human competencies in tasks like pattern recognition, playing board games, and outputting text-based expert information in multiple domains, especially with LLMs. Some people are of the view that, as GPT-4 is capable of solving

"novel and difficult tasks that span mathematics, coding, vision, medicine, law, psychology and more, without needing any special prompting," (Bubeck et al., 2023), it already meets the hallmarks of artificial general intelligence (AGI). Jones and Bergen (2025) make similar claims for GPT-4.5, which aces the Turing test with a "win rate" of 73% of convincing human judges that the AI is human, while humans struggle to do so themselves. However, there has been pushback on AI acing Turing tests as being insufficient, or even a case of misdirected evidence of intelligence. Mitchell (2024) claims Turing tests suffer from moving goal posts due to "our shifting conceptions of intelligence." The capacity of machines to hold fluent conversations in natural language that Turing proposed in 1950 is no longer considered to be evidence of general intelligence. Whether feats of GPT-4 LLMs qualify to be on par with human cognition, which marks an apogee for general intelligence, is part of ongoing debates (see Goertzel, 2014; Zimmermann, 2024).

Many have characterized human-level intelligence as having broad-ranging, adaptive powers that can respond to changing external environments by selecting goals and the means to achieve them by including novel solutions. For instance, having given a long list of characteristics of human-level intelligence, which includes self–other awareness and self-control, Goertzel (2014) requires AGI to have "general scope and is good at generalization across various goals and contexts." With regard to novel solutions, to date, the open-ended adaptive capacity of humans produces what Dawkins (1989) calls extended phenotypes or artifacts outside of ourselves, rather than following a trans-human agenda with genomic enhancements.

It has become commonplace to state that intelligence is what mediates the goals–means nexus and is characteristic of goal-directed agents (Russell and Norvig, 2003). The pushback on GPT LLMs on having a Q&A format, in which the AI does not learn anything, due to insufficient experientially driven data from the environment actively elicited by the agent, has been raised by Silver and Sutton (2024). Their vision of the next stage of AI agents is of those that are autonomous in their selection of goals and means, capable of self-learning from a continuous stream of experientially driven feedback governed by reward maximization. However, we have here the infamous proclamation of Captain Ahab in Moby Dick, "All my means are sane, my motive and my object are mad." This calls into question what "sane" goals are, if the only hallmark of rationality qua intelligence, in extant decision sciences (see Markose, 2024; Silver et al., 2021), is the reward maximization calculus of efficiency in the service of an objective.

In recent discussions, the AI control problem or the misalignment problem (Bostrom, 2014; Russell, 2019; Ngo et al., 2023; Hinton, 2023) has been recognized when AI systems are autonomous and evolve malign behaviors that may not align with human values and can evade human control. The Ngo et al. (2023) description of AI agents that use deception and power-seeking strategies to pursue misaligned goals underscores this as a perennial problem of political economy that is not unique to AI. At least since the Hobbesian thesis on the struggle for power and resources, it has been recognized that there is an existential threat to life and society when an agent with unbridled adaptive intelligence is free to set its own goals and encounters other similarly intelligent agents with their goals. The problem of adversarial and conflicting goals is writ large. The extant computational environment is swarming with sniffers, snipers, deep fakes, and computer viruses. In all cases, though these bots are installed by humans, they can operate with various degrees of autonomy to deceive, defraud, bring down software systems, and, in the case of killer bots and drones, physically decapitate humans and destroy their digital and material possessions. Generative adversarial networks (GANs), for instance, can program bots to resist detection by making deep fakes of themselves.

In this note, I aim to throw new light on three aspects of the control or misalignment problem of AGI. For this, I will draw on advances in gene, neuro, and computer sciences, especially cryptography, on how to protect purposeful software systems that can lose autonomy when malware agents can hack and hijack host codes to do their bidding. What is interesting is that autonomy and selfhood, often considered to be vestiges of liberal democracy, are part of the unique information processing of a code-based system of life that has maintained the unbroken chain of life while permitting evolutionary change.

The first step is to refer to the above discussions, the provenance of general intelligence as a means of maintaining homeostasis of life (Friston, 2010; Friston, 2013). In other words, the fundamental alignment of general intelligence is in the service of life itself and not any narrow objective. However, I will replace the Friston et al. Free Energy principle for self-organization of life's homeostasis in terms of minimizing the degrading forces of entropy and disorder with a code-based explanation for general intelligence. I will elaborate on how the digital socio-economic world driven by AI has parallels with what I call genomic intelligence (Markose, 2024), which accords with the Walker and Davis (2013) epigram on the "algorithmic take-over" of biology with digitization of inheritable information encoded in a near-universal alphabet (A, T, C, G/U) in the genome.

Second, I will introduce the reader to the adversarial digital game, coextensive with life itself as the fundamental source of misalignment, that was brought to my attention by the game theorist Ken Binmore (see Markose, 2021c). Binmore (1987) raised the specter of Gödel's Liar, qua digital adversary, who will negate what can be predicted. Binmore uses Gödel's Liar to highlight the flaw of extant Game Theory: by confining the best response to a given action set, Game Theory not only guarantees that determinism will be punished by the Liar but also precludes novelty and surprises in the Nash equilibrium of a game. Markose (2017) produces a Nash equilibrium of a game with Gödel's Liar, which, from logical necessity, produces novel syntactic objects outside listable sets. To date, complexity, evolvability, novelty production, and "thinking outside the box" in biology and humans have, for the most part, relied on models of randomness or on statistical white noise error terms (Markose, 2024; Markose, 2021b). This is despite the long-standing type IV dynamics in the Wolfram–Chomsky schema, based on foundational mathematics of Gödel, Turing, and Post (G–T–P) aka recursion function theory (RFT), that only code-based computational systems that can embrace self-referential recursive structures of the Gödel incompleteness theorems (GITs) can produce novelty (see Prokopenko et al., 2019; 2025). Given that for some 90 years there has been little evidence that GITs and the capstone construction in the form of the Gödel sentence has relevance to any real world phenomena, in Section 2, I will unpack some of the recent evidence of how such self-referential intelligence was acquired for complexification

over the course of evolution of multicellular eukaryote life (Markose, 2022).

Third, a major development of the 21st-century digital age, which has a bearing on the misalignment problem, is the astounding invention of the blockchain distributed ledger technology (BCDL). This was first presented in the anarchic agenda of Bitcoin by pseudonymous Nakamoto (2008) to resist centralized state control of monetary systems. BCDL permits decentralized software-based record keeping of actions of multiple agents, in which the fidelity of extant digital records is maintained by a hashing solution to a cryptographic puzzle. This also makes it difficult for malign activity regarding new software additions by a subset of agents. Abramov et al. (2021), Markose (2021a), and Markose (2022) have been the first to point out that the genome is a blockchain. However, while Abramov et al. (2021) utilize the consensus mechanism well known in manmade blockchain (Hussein et al., 2023). Markose (2022) indicates that the genomic blockchain relies on a self-referential hashing solution using the Gödel sentence, which permits biotic elements to self-report that they are under attack. The immutability of protein coding blocks of life for 3.7 billion years, associated with Crick's notion of a "frozen accident" while novelty is added, in the 21st century, can be identified as part of a unique self-referential BCDL embodied in the organism that secures alignment with the homeostasis of life. In any case, there is a growing understanding that unless software systems are embedded in a BCDL, they will be doomed to failure by optimization of narrow objectives, as in the Bostrom (2014) paperclip apocalypse, or hacked to pieces due to the sitting duck problem (see Nabben, 2021; Heaven, 2019).

2 Staples of G–T–P/RFT, genomic intelligence, and homeostasis of life

Until recently, there has been little evidence of how the staples of RFT and Gödel (1931) relate to genomic systems, let alone to BCDLs. This section will unpack the breakthroughs on the evidence that RFT staples are ubiquitous in the self-referential genomic intelligence of eukaryotes.

2.1 Unique digital identifiers and hashes in biology

First, which is also a major ingredient of BCDLs for malware detection, is the feature of unique digital identifiers pioneered by Gödel (1931), called Gödel numbers (g.n.) or indexes, whereby a finite string of letters maps to a unique integer. The hash compresses variable-length strings to a fixed length, and any change in input strings will alter the hash. There is now extensive evidence of bio-peptide and other unique identifiers, including "zip codes" for cellular signal processing, as discovered in the Nobel prize-winning work of Blobel (1999). It appears that all signaling in bio-ICT relies on peptide identifiers from transcription factors in gene expression to neuron-neuron links. As in the design of BCDLs that all nodes of the distributed system have the same information, more than 30 trillion cells in a human have the same DNA, with some exceptions of mosaicism. There is evidence (see Brickner et al., 2012) that subnetworks of gene regulatory networks have characteristic identifiable binding motifs in transcription factors and their binding sites for associated gene expression for temporal and specialized cell development in tissues. We will denote by g, $g \in G$, the DNA instructions that lead to gene expression of specific somatic and phenotype developments of the organism, where G is the set of expressed genes.

2.2 Self-reference and diagonal self-assembly machines in biology

In RFT, using epithets from Hofstader (1999), we have self-reference (Self-Ref) or diagonal operators typically stated as a program, g, that builds a machine that runs g and halts (denoted as $\Phi_g(g)\downarrow$). Gershenfeld (2012) and Gershenfeld et al. (2017) give the remarkable insight that what 21st-century digital fabrication aims to do, which is described as end-to-end code-based 3-D self-assembly of digitized materials, is something biology solved 3.7 billion years ago. The self-assembly programs of biology are associated with the ribosome and other transcriptase machinery that implement gene expression for the morphological, somatic identity, and regulatory control of the organism.

The breakthrough on the significance of this staple of self-referential/diagonal operator in RFT found in textbooks like Rogers (1967) and Cutland (1980) for biological self-assembly is given in Panel A of Figure 1. Following the set theoretic proof of GITs in Post (1944), Cutland (1980), and Smullyan (1961), $g \in G$, that determine selfhood of the organism can be considered the theorems for the organism and G–T–P information processing, and alignment to the homeostasis of life is stringently governed by the principle of consistency.

2.3 Offline self-representation (Self-Rep) or mirror mappings of online self-assembly machine executions

A major breakthrough here is the evidence Markose (2021a), Markose (2022) found for Self-Rep mirror structures of the adaptive immune system (AIS), approximately 500 mya post jawed fish, which is not present in prokaryotes. The major histocompatibility complex (MHC1) receptors of the thymus are found to record 85% of expressed genes relating to the 3D self-assembly of the morphology and somatic identity of the organism. This is shown, respectively, in the left (offline recording) and right (online self-assembly Self-Ref machine execution) sides of Panel A of Figure 1. For good reason, these self-repped gene codes in the thymus have been called the *Thymic Self*, Sánchez-Ramón and Faure (2020) or "the science of self" Greenen (2021). As is well known, the Self-Rep in AIS thymic receptors is primarily to identify the hostile other, viz., negation function operators of non-self-antigens, as will be discussed below. Indeed, Miller et al. (2019) wax lyrical: "As self-referential cognition is demonstrated by all living organisms, life can be equated with the sustenance of cellular homeostasis in the continuous defense of 'self'." This is remarkable in that Miller et. al. give centrality to self-referential information processing in genomic systems specifically to detect and mitigate adversarial

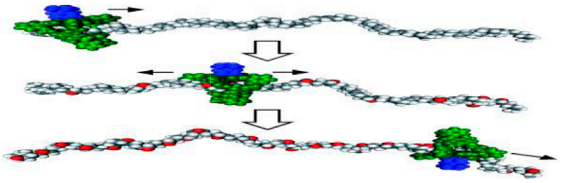
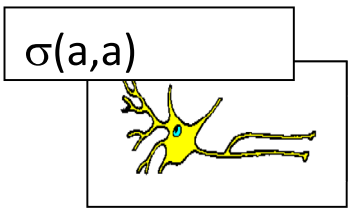
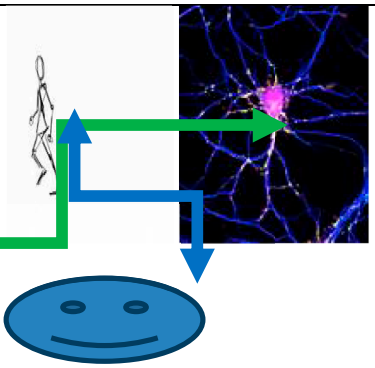
Offline Recording/Simulation *(LHS) Self-Rep* *Notation* $\phi_{\sigma(g,g)}(s)$	*Online Machine Execution* *(RHS) Self-Ref /Diagonal Machine* *Notation:* $\phi_{\phi_g(g)}(s)=q,$
Panel A (Left) **Medullary Thymus Epithelial cells (m-TECs) Mirror Mapping** of tissue specific genes/peptides of ~85% of genome that is expressed	**Panel A (Right)** **Baseline Machine Execution of protein Coding and non-coding Genes modelled as a self-assembly of digitized biotic elements of gene codes in set G**
Baseline : With no malware σ(g,g)	Ribosomal and non-coding Transcription machine executions on *g* codes which halt: $\phi_g(g) \downarrow$ *for all* $g \in$ G 
Panel B (Left) Mirror Neurons (Yellow) For Social Cognition: Fires with self action and when self 'sees' others' actions **For action prediction with self and other**	**Panel B (Right) Canonical Neurons (Pink on black) Firing Based on Motor Cortex (eg'self'-walking, Green Arrows)and Sensory/Visual Cortex as observer by self of 'other' walking (Blue Arrows)**
Discovered by Parma Group '90s (Gallese, Fadiga, Rizollatti, Di Pellegrino, Fogassi) σ(a,a) 	

FIGURE 1
Gödel meta-representation (Rogers, 1967) and mirror systems in immuno-cognitive systems. Note: *Offline* mirror systems in the medulla thymus (Panel A, Left) and *Offline* cognitive mirror neuron system (Panel B, Left) and a respective bijective map of *Online* gene transcription (Panel A, Right) and *Online* action execution in the motor–sensory cortex (Panel B, Right).

changes to self-codes but make no reference to the RFT staple of the Gödel representation theorem from Rogers (1967), which is exactly depicted in Figure 1. As a result, Miller et al. (2019) is a compendium of analogies and possible inaccuracies but lacks RFT or a code-based explanation of how self-referential structures implement the defense of self-codes.

An even larger literature in neuroscience on mirror mappings has followed important discoveries of the Parma Group (Fadiga et al., 1995; Gallese et al., 1996; Rizzolatti et al., 1996; Gallese and Sinigaglia, 2011; Gallese, 2009) of a mirror neuron system (MNS) in the primate brain where embodied self-action codes from the sensory-motor cortex are mapped *offline* and reused

to make action inference in conspecifics and help facilitate complex self–other interactions (see also Markose et al. (2025)). However, despite the central role assigned to self-reference for the sentient self in advanced organisms (Gardenfors 2003; Northoff et al., 2006; Newen, 2018; Miller et al., 2019, etc.), only Tsuda (2014), Markose (2017), Markose (2021b), and Markose (2022) have noted how the evolutionary development of Self-Rep offline mirror structures is necessary for biotic elements to make statements about themselves. Tsuda (2014) makes an explicit observation that unless the two-step mirror Self-Rep recursive structures are in place, the mapping between the online machine execution codes and the offline recording of the same shown in Figure 1 using the σ (x,x) 2 – place, the Gödel substitution function, it is unlikely that statements about self can be made, let alone about the other.

2.4 How can changes to genomic self-codes be detected, specifically those brought about by a digital adversary?

Here, the breakthrough in gene science, which debunks the idea that the primary source of evolutionary changes arises from random transcription/replication errors, follows the epochal discovery by Nobel Laureate Barbara McClintock (1984) of transposable elements (TEs) of viral origin. TEs that conduct cut-paste (transposons) and copy-paste (retrotransposons) give a code-based explanation for genomic changes. TEs, which account for some 45% of the genome, have been found to engineer genomic evolvability, brain plasticity, and novel phenotypes primarily in eukaryotes (Fedoroff, 2012). This underscores the truism that only software can change software and also sheds light on the double-edged sword of viral software. It can benignly copy and paste as in replication, which entails a simple sliver of code, as shown in one of the earliest accounts of code biology by Kauffman (2015). However, malign viral hacking, done externally by bio-malware and or internally by TEs to gene expression itself, forms the Achilles heel of genomic digital systems.

Hence, here we have the model for the self-referential detection of Gödel's Liar. This entails the adaptive immune system (AIS) (Flajnik and Kasahara, 2009) in the T-cell receptors that "simulate" the application of negation software functions, $f^{\neg}$, qua virus (hacker) on self-repped gene codes. The breakthrough here is to see that an RFT generalization of Gödel (1931) using Roger's Fixed Point Theorem (Rogers,1967) is needed for the counterparts in the periphery of the self-repped gene codes in the T-cell receptors to self-report when software changes to self-codes are brought about by novel non-self-antigens in real time. The latter are an uncountable infinity.

The AIS implements "out of the box" astronomic anticipative search for novel non-self-antigens necessary for novel antibody production and cognition in humans, manifesting unbounded proteanism for novel extended phenotypes (Dawkins, 1989) in the form of artifacts outside of ourselves. This facility, first found in the AIS, relies on the recombination activator genes (RAG 1 and 2) and also in the human brain for neural receptor diversity (Muotri et al., 2009; Kaesar and Chun, 2020; Peña de Ortiz and Arshavsk, 2001), which runs into orders of magnitude of 10^{20}–10^{30} (Kapitonov and Jurka, 2005) that exceed the pre-scripted germline of the genome size many times over. Likewise, detection of negation of what is predicted in the human mirror neuron system found in neuroscience experiments by Scott Kelso and co-authors (Tognoli et al., 2007) gives evidence for perception of deceit and complex counterfactuals in the Theory of Mind in social cognition.

The Rogers (1967) fixed point indexes of the Second Recursion Theorem for yet-to-happen $f^{\neg}$ attacks by the non-self-antigens are generated in the AIS in a most ingenious fashion: a large number of codes/indexes purported to be of different $f^{\neg}$ on each self-repped g are generated in the T-cell receptors. This is the most spectacular case of predictive coding. Suppose that the g.n for the tuple $\{ f^{\neg}, g \}$ specifying that $f^{\neg}$ has attacked g, is denoted by $g^{\neg}$. When the attack by $f^{\neg}$ takes place in real time in the periphery involving g, the experientially driven peripheral MHC1 receptor mediated by interferon gamma must record this. If this "syncs" with the one that was speculatively generated in the thymic MHC1 receptors, two parts of the fixed point come together to construct a genomic Gödel sentence, which will now have a fixed-point index of $\sigma (g^{\neg}, g^{\neg})$. At this point, g self-reports that it is under attack.

The index $\sigma (g^{\neg}, g^{\neg})$ of the Gödel sentence effectively signals the hash for an untenable state of 0 = 1 produced by the fixed point of a $f^{\neg}$ negation function of self-codes (see Kauffman, 2023). Such syntactic objects, $\sigma (g^{\neg}, g^{\neg})$ at the point at which it is recursively generated, are undecidable in that they lie outside of listable sets arising from the mapped self-repped expressed gene codes that are the theorems for the organism and the list of indexes for known non-theorems. Such indexes $\sigma (g^{\neg}, g^{\neg})$ of Gödel sentences have recently been identified by Markose (2017), Markose (2021a), and Markose (2022) as a precursor for endogenous novelty production in genomic systems. Indeed, it is a testable hypothesis that it is the inability of the peripheral MHC1 receptor to update the index to $\sigma (g^{\neg}, g^{\neg})$ when the $f^{\neg +}$attacks g, typically due to faulty interferon gamma mediation, that causes AIS to fail to generate novel antibodies (Markose, 2021a). In RFT, the productive set of Post (1944) provides the unique recursive construction of the blockchain of fixed point indexes $\sigma (g^{\neg}, g^{\neg})$ for the novel non-self-antigens and the novel antibodies thereof. This takes on the structure of an arms race, which is somatic and *extraneous* to the germline; hence, this exercise is geared to conserve the genome rather than to improve it.

Schmidhuber (2006) and more recent articles (Zhang et al., 2025) have depicted Gödel machines and Darwin Gödel machines, respectively, to show how self-referential mappings can lead to self-improving machines that can rewrite their own codes. It is important to note here that the precise implementation of the structures of Gödel incompleteness as found in the adaptive immune system, which involves the detection of novel negation functions of adversarial agents and their fixed-point indexes as in the Gödel sentence, the novel antibody production that follows does not lead to self-improvement in the germline. Instead, the self-referential recursive structures are geared toward conserving self-codes against adversaries, and the arms race in novelty is to improve defenses and maintain autonomy of the organism against prolific digital adversaries.

It is conjectured that an identical RFT machinery is involved in the self–other nexus in both the AIS and MNS. What evidence is there for this? In a knockout of interferon gamma in the Jonathan Kipnis Group experiment on rats (Filiano et al., 2016), it was found that the rats lost immune capabilities as well as their social cognition of recognizing another rat. Kipnis et al. give an Evo-Devo

explanation that evolution has taught rats to socially isolate when their immune system is compromised. My code-based explanation (see Markose, 2021a) is that the same self-referential recursive structures are in place *both* for the AIS as well as in the brain MNS for self–other cognition and hence when the interferon gamma mediator, especially in the *peripheral* MHC1 receptor is knocked out, the circuitry for the fixed point generation needed for predictive coding for non-self–other misfires and self becomes blind to the other. It is conjectured that this is how the rats in the Filiano et al. (2016) experiment lost their immune capabilities *and* their capacity for social cognition of another rat.

3 Concluding remarks

In conclusion, genomic intelligence in vertebrates that has reached its pinnacle in humans is highly empathic as the conspecific/other is the projection of self; greatly Machiavellian having co-evolved from adversarial viral agents; geared toward unbounded proteanism from the get-go starting with transposon-based diversity creation by recombination activation genes (RAG) in the immune system and brain; and stringently self-regulated by a self-referential block chain distributed ledger (BCDL) driven by the principle of autonomy of the life of the organism and an agenda to be hack free.

It is a matter of incredulity that some 90 years have passed since Gödel (1931), for evidence to be found that the RFT staples of Self-Ref and Self-Rep and the Gödel sentence are ubiquitous in biology and genomic intelligence. Several factors can be adduced for the lack of precise computational modeling of self-reference in the context of general intelligence. Even those who espouse that code-based operations are relevant in cognition, such as in the Computational Theory of Mind (see Rescorla, 2020), never mention Self-Ref, self-assembly machines, Self-Rep mirror systems, or computational fixed-point indexes and, of course, the role of the Gödel sentence. There is a strong anti-machine view that claims that biology is a non-digital "natural" process that is creative in some vitalistic way. This view overlooks the fact that in nature, only biology, with the encoded basis of the genome, and the extended phenotypes of humans who have built computers, manifest software-based digital systems.

Two canards are associated with the Gödel incompleteness theorems (GITs) that seem to propagate anti-machine vitalistic beliefs about life and intelligence. These have posed a stumbling block to the necessary breakthroughs on code-based explanations for genomic information processing. The canards are that the GIT proves that human cognition is not computational and self-reference leads to paradox (see Battaglia et al., 2025). In the Gödelian setting, as unlike the Cretan Liar paradox *This is False*, Gödel's painstaking two-step process of Self-Ref and Self-Rep, found in Rogers (1967) generalizations thereof, on how statements about self and other appear to be made in the immune-cognitive systems (see Figure 1; Markose, 2021b), there are no paradoxes. Furthermore, influential commentators like Roger Penrose have used the GIT to conclude that human cognition can outstrip what Turing machines can do. As Rescorla (2020) says, "It may turn out that certain human mental capacities outstrip Turing-computability, but Gödel's incompleteness theorems provide no reason to anticipate that outcome." The work of LaForte et al. (1998), *On Why Gödel's Theorem Cannot Refute Computationalism,* and others has provided push back on such flawed anti-machine views on biology and human cognition.

Section 2 gives an account of how G–T–P based immune-cognitive systems may be conducting self-referential information processing. As first noted by Tsuda (2014), it is unlikely that any statements regarding self or the other can be made by humans without the two-step Self-Ref and Self-Rep recursive information processing structures having evolved. Furthermore, the recursive generation of an index of the Gödel sentence should demystify what it is, a hash representing "0 = 1," viz., non-theoremhood or misalignment, and how it signifies a novel object as a constructive "witness" for proof of incompleteness. In view of this, the following statement is misconstrued: "The paradox of a brain trying to study itself presents a conundrum, raising questions about self-reference, consciousness, psychiatric disorders, and the boundaries of scientific inquiry" (Battaglia et al., 2025). Likewise, in the absence of the precise recursive function structures of Self-Ref and Self-Rep necessary to identify software changes to self-codes, the important discovery by the Parma Group of the mirror neuron system in the brain has been stymied by hype and inaccuracies.

In the influential non-code-based Free Energy principle explanation for general intelligence involved in the homeostasis of life (Friston, 2013), Friston does not fall into the trap of the mainstream optimization framework, which effectively constrains search to under the lamp post and cannot produce novelty. Schwartenbeck et. al (2013) state that search for novel solutions and "explorative behavior is not just in accordance with the principle of free energy minimization but is, in fact, mandated by it." However, from the vantage of the discussion here, it seems that there has been insufficient consideration by Friston of the regulatory framework of maintaining homeostasis of life's vital signs within feasible physical/analog states, viz., minimizing "surprisals," when this is under the aegis of smart algorithmic controls. The latter must contend with software-related data security breaches from bio-malware or adversarial digital agents.

What has been overlooked is that a large part of homeostasis in formalistic code-based self-assembly systems of life involves the complexification of phenotype with dynamic adversarial digital game structures that must embrace an arms race in novelty and surprises in order to avoid threats to autonomy from adversarial agents that can hack gene codes. This is a problem that genomic intelligence appears to have solved. AI, in contrast, has ignored this self-referential design for data integrity for autonomous existence that can vitiate what is called the "sitting duck" problem (Heaven, 2019). Furthermore, extant decision sciences are devoid of any epistemic structures for novelty production and complexification (Markose, 2024). As noted, there is a considerable difference between the Gödel self-reference models for novelty production by Markose (2021c), Markose (2022) and those of Schmidhuber (2006) and Zhang et al. (2025). I underscore the formal system premise of consistency and theoremhood for providing the stringent selection mechanism for what novelty is permitted in genomic systems and do not use the language of optimal self-improvement or reward frameworks as do these and other authors.

In closing, it is my view that the biological immuno-cognitive model of the self-referential genomic BCDL with Gödel sentence

hashes has far-reaching implications for understanding the full gamut of self–other pathologies, gene regulatory networks that must deal with malign transposable element activity, and more robust design solutions for sustainable AGI.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author.

Author contributions

SM: Writing – review and editing, Conceptualization, Investigation, Formal Analysis, Writing – original draft, Visualization.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

I am grateful for discussions and inputs from Mikhail Prokopenko, Karl Friston, Ken Binmore, Georg Northoff, Louis Kauffman, Patrick Hayes, Neil Gershenfeld, Gordana Dodig Crnkovic, and Pedro C. Marijuán. I remain responsible for any errors.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

References

Abramov, O., Bebell, K. L., and Mojzsis, S. J. (2021). Emergent bioanalogous properties of blockchain-based distributed systems. *Orig. Life Evol. Biosph.* 51, 131–165. doi:10.1007/s11084-021-09608-1

Battaglia, S., Servajean, P., and Friston, K. (2025). The paradox of the self-studying brain. *Phys. Life Rev.* 52, 197–204. doi:10.1016/j.plrev.2024.12.009

Binmore, K. (1987). Modeling rational players: *part I. J. Econ. Philosophy* 3, 179–214. doi:10.1017/s0266267100002893

Blobel, G. (1999). The nobel prize in physiology or medicine 1999. Available online at: https://www.nobelprize.org/prizes/medicine/1999/summary/ (Accessed March 2, 2025).

Bostrom, N. (2014). *Superintelligence: paths, dangers, strategies*. OUP.

Brickner, D. G., Ahmed, S., Meldi, L., Thompson, A., Light, W., Young, M., et al. (2012). Transcription factor binding to a DNA zip code controls interchromosomal clustering at the nuclear periphery. *Dev. Cell.* 22 (6), 1234–1246. doi:10.1016/j.devcel.2012.03.012

Bubeck, S., Chandrasekaran, V., Ronen Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., et al. (2023). Sparks of artificial general intelligence: early experiments with GPT-4. *arXiv:2303.12712 [cs.CL]*.

Cutland, N. J. (1980). *Computability: an introduction to recursive function theory*. Cambridge University Press.

Dawkins, R. (1989). *The extended phenotype. Oxford*. Oxford University Press.

Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiology* 73, 2608–2611. doi:10.1152/jn.1995.73.6.2608

Fedoroff, N. V. (2012). Transposable elements, epigenetics and genome evolution. *Science* 338, 758–767. doi:10.1126/science.338.6108.758

Filiano, A., Xu, Y., Tustison, N., Marsh, R. L., Baker, W., Smirnov, I., et al. (2016). Unexpected role of interferon-γ in regulating neuronal connectivity and social behaviour. *Nature* 535, 425–429. doi:10.1038/nature18626

Flajnik, M. F., and Kasahara, M. (2009). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* 11, 47–59. doi:10.1038/nrg2703

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi:10.1038/nrn2787

Friston, K. (2013). Life as we know it. *J. R. Soc. Interface* 10, 20130475. doi:10.1098/rsif.2013.0475

Gallese, V. (2009). Mirror neurons, embodied simulation, and the neural basis of social identification. *Psychoanal. Dialogues* 19 (5), 519–536. doi:10.1080/10481880903231910

Gallese, V., and Sinigaglia, C. (2011). What is so special about embodied simulation? *Trends Cognitive Sci.* 15 (11), 512–519. doi:10.1016/j.tics.2011.09.003

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119 (2), 593–609. doi:10.1093/brain/119.2.593

Gardenfors, P. (2003). *How homo became sapiens: on the evolution of thinking*. Oxford University Press.

Gershenfeld, N. (2012). "How to make anything: the digital fabrication revolution," in *Fourth industrial revolution, the davos reader*. Editor G. Rose

Gershenfeld, N., Gershenfeld, A., and Cutcher-Gershenfeld, J. (2017). *Designing reality: how to survive and thrive in the third digital revolution*. New York: Basic Books, Hachette Book Group.

Gödel, K. (1931). "On formally undecidable propositions of *principia mathematica* and related systems," in *Translation in English in gödel's theorem in focus*. Editor S. G. Shanker (Croom Helm).

Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. *J. Artif. General Intell.* 5 (1), 1–48. doi:10.2478/jagi-2014-0001

Greenen, V. (2021). The thymus and the science of self. *Semin. Immunopathol.* 2021, 1–10.

Heaven, D. (2019). Why deep-learning AIs are so easy to fool: artificial-intelligence researchers are trying to fix the flaws of neural networks. *Nat. News Feature*. Available online at: https://www.nature.com/articles/d41586-019-03013-5 (Accessed March 2, 2025).

Hinton, G. (2023). The godfather of A.I.' warns of 'nightmare scenario' where artificial intelligence begins to seek power. *Fortune*.

Hofstader, D. (1999). *Gödel, escher, bach: an eternal golden braid*. Basic Books.

Hussein, Z., Salama, M. A., and El-Rahman, S. A. (2023). Evolution of blockchain consensus algorithms: a review on the latest milestones of blockchain consensus algorithms. *Cybersecurity* 6, 30. doi:10.1186/s42400-023-00163-y

Jones, C., and Bergen, B. (2025). *Large language models pass the turing test*. doi:10.48550/arXiv.2503.23674

Kaesar, G., and Chun, J. (2020). Brain cell somatic gene recombination and its phylogenetic foundations. *J. Biol. Chem.* 295 (36), 12786–12795. doi:10.1074/jbc.rev120.009192

Kapitonov, V. V., and Jurka, J. (2005). RAG1 core and V(D)J recombination signal sequences were derived from transib transposons. *PLoS Biol.* 3, e181. doi:10.1371/journal.pbio.0030181

Kauffman, L. (2015). Self-reference, biologic and the structure of reproduction. *Prog. Biophys. Mol. Biol.* 119 (3), 382–409. doi:10.1016/j.pbiomolbio.2015.06.015

Kauffman, L. (2023). Autopoiesis and eigenform. *Computation* 11 (12), 247. doi:10.3390/computation11120247

LaForte, G., Hayes, P., and Ford, K. (1998). Why Gödel's theorem cannot refute computationalism. *Artif. Intell.* 104 (Issues 1–2), 265–286. doi:10.1016/S0004-3702(98)00052-6

Markose, S. M. (2017). Complex type 4 structure changing dynamics of digital agents: nash equilibria of a game with arms race in innovations. *J. Dyn. Games* 4 (3), 255–284. doi:10.3934/jdg.2017015

Markose, S. M. (2021a). Genomic Intelligence as Über Bio-Cybersecurity: the Gödel Sentence in Immuno-Cognitive Systems. *Entropy* 23, 405. doi:10.3390/e23040405

Markose, S. M. (2021b). "Novelty production and evolvability in digital genomic agents: logical foundations and policy design implications of complex adaptive systems," in *Complex systems in the social and behavioral sciences: theory, method and application*. Editors E. Elliot, and L. Douglas Kiel (Ann Arbor, MI, USA: Michigan University Press).

Markose, S. M. (2021c). How we became smart-a journey of discovery through the world of game theory and genomic intelligence. *Essex Blog*. Available online at: https://www.essex.ac.uk/blog/posts/2021/10/26/how-we-became-smart (Accessed on March 2, 2025).

Markose, S. M. (2022). Complexification of eukaryote phenotype: adaptive immuno-cognitive systems as unique Gödelian blockchain distributed ledger. *Biosystems* 220, 104718. doi:10.1016/j.biosystems.2022.104718

Markose, S. M. (2024). "Digital foundations of evolvable genomic intelligence and human proteanism: complexity with novelty production beyond bounded rationality," in *Routledge International Handbook of Complexity Economics*. Editors P. Chen, W. Elsner, and A. Pyka, 528–550.

Markose, S. M., Friston, K., Northoff, G., Cross, E., and Prescott, T. (2025). Frontiers special topic: narrow and general intelligence: embodied, self-referential social cognition and novelty production in humans, AI and robots. Available online at: https://www.frontiersin.org/research-topics/25980/narrow-and-general-intelligence-embodied-self-referential-social-cognition-and-novelty-production-in-humans-ai-and-robots.

McClintock, B. (1984). The significance of responses of the genome to challenge. *Science* 226 (4676), 792–801. doi:10.1126/science.15739260

Miller, W., Torday, J., and Baluška, F. (2019). Biological evolution as defense of 'self. *Prog. Biophysicsand Mol. Biol.* 142, 54–74. doi:10.1016/j.pbiomolbio.2018.10.002

Mitchell, M. (2024). The turing test and our shifting conceptions of intelligence. *Science* 385 (6710), eadq9356. doi:10.1126/science.adq9356

Muotri, A., Zhao, C., Marchetto, M., and Gage, F. (2009). Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus* 19, 1002–1007. doi:10.1002/hipo.20564

Nabben, K. (2021). Decentralised autonomous organisations (DAOs) as data trusts: a general-purpose data governance framework for decentralised data ownership, storage, and utilisation (december 20, 2021). *SSRN*. doi:10.2139/ssrn.4009205

Nakamoto, S. (2008). Bitcoin: a peer-to-peer electronic cash system. *Decentralized Bus. Rev.*, 21260.

Newen, A. (2018). The embodied self, the pattern theory of self, and the predictive mind. *Front. Psychol.* 2270. doi:10.3389/fpsyg.2018.02270

Ngo, R., Chan, L., and Mindermann, S. (2023). The alignment problem from a deep learning perspective. *arXiv:2209.00626 [cs.AI]*. doi:10.48550/arXiv.2209.00626

Northoff, G., Heinzel, A., Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage* 31 (1), 440–457. ISSN 1053-8119. doi:10.1016/j.neuroimage.2005.12.002

Peña de Ortiz, S., and Arshavsky, Y. (2001). *DNA recombination as a possible mechanism in declarative memory: a hypothesis journal of neuroscience*. Wiley Online Library.

Post, E. (1944). Recursively enumerable sets of positive integers and their decision problems. *Bull. Am. Math. Soc.* 50, 284–316. doi:10.1090/s0002-9904-1944-08111-1

Prokopenko, M., Harré, M., Lizier, J., Boschetti, F., Peppas, P., and Kauffman, S. (2019). Self-referential basis of undecidable dynamics: from the liar paradox and the halting problem to the edge of chaos. *Phys. Life Rev.* 31, 134–156. doi:10.1016/j.plrev.2018.12.003

Prokopenko, M., Davies, P. C. W., Harré, M., Heisler, M. G., Kuncic, Z., Lewis, G. F., et al. (2025). Biological arrow of time: emergence of tangled information hierarchies and self-modelling dynamics. *J. Phys. Complex.* 6, 015006. doi:10.1088/2632-072X/ad9cdc

Rescorla, M. (2020). *The computational theory of mind, the stanford encyclopedia of philosophy (fall 2020 edition)*. Editor E. N. Zalta Available online at: https://plato.stanford.edu/archives/fall2020/entries/computational-mind/.

Rizzolatti, G., Fadiga, L., Gallese, V., and Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Res.* 3, 131–141. doi:10.1016/0926-6410(95)00038-0

Rogers, H. (1967). *Theory of recursive functions and effective computability*. McGraw-Hill.

Russell, S. (2019). *Human compatible: artificial intelligence and the problem of control*. Penguin Random House.

Russell, S. J., and Norvig, P. (2003). *Artificial intelligence a modern approach (2nd edn)*. Upper Saddle River, NJ: Prentice Hall.

Sánchez-Ramón, S., and Faure, F. (2020). Self and the brain: the immune metaphor. *Front. Psychiatry* 2020 (11), 540676. doi:10.3389/fpsyt.2020.540676

Schmidhuber, J. (2006). Gödel machines: self-referential universal problem solvers making provably optimal self-improvements. *arXiv:cs/0309048 [cs.LO]*. doi:10.48550/arXiv.cs/0309048

Schwartenbeck, P., FitzGerald, T., Dolan, R., and Friston, K. (2013). Exploration, novelty, Surprise, and free energy minimization. *Front. Psychol.* 4, 710. doi:10.3389/fpsyg.2013.00710

Silver, D., and Sutton, R. (2024). *Welcome to the era of experience, forthcoming designing an intelligence*. published by MIT Press.

Silver, D., Singh, S., Precup, D., and Sutton, R. S. (2021). Reward is enough. *Artif. Intell.* 299, 103535. doi:10.1016/j.artint.2021.103535

Smullyan, R. (1961). *Theory of formal systems*. Princeton University Press.

Tognoli, E., Lagarde, J., DeGuzman, G., and Kelso, S. (2007). The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. U. S. A.* 104 (19), 8190–8195. doi:10.1073/pnas.0611453104

Tsuda, I. (2014). "Logic dynamics for deductive inference its stability and neural basis," in *Chapter 17 In, chaos, information processing and paradoxical games: the legacy of john S nicolis*. Editors N. Gregoire, and B. Vasileios (World Scientific Publishing Co. Pte.Ltd.).

Walker, S. I., and Davies, P. C. W. (2013). The algorithmic origins of life. *J. R. Soc. Interface* 10, 20120869. doi:10.1098/rsif.2012.0869

Zhang, J., Hu, S., Lu, C., Lange, R., and Clune, J. (2025). Darwin Godel machine: open-ended evolution of self-improving agents. *arXiv:2505.22954 [cs.AI]* 2. doi:10.70777/si.v2i3.15063

Zimmermann, G. (2024). What makes systems intelligent. *Discov. Psychol.* 4, 127. doi:10.1007/s44202-024-00245-z

frontiers | Frontiers in Robotics and AI

TYPE Original Research
PUBLISHED 22 September 2025
DOI 10.3389/frobt.2025.1586033

OPEN ACCESS

EDITED BY
Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY
Daniele Meli,
University of Verona, Italy
Asieh Abolpour Mofrad,
University of Bergen, Norway

*CORRESPONDENCE
Robert Johansson,
robert.johansson@psychology.su.se

RECEIVED 01 March 2025
ACCEPTED 13 August 2025
PUBLISHED 22 September 2025

CITATION
Johansson R (2025) Modeling arbitrarily applicable relational responding with the non-axiomatic reasoning system: a Machine Psychology approach.
Front. Robot. AI 12:1586033.
doi: 10.3389/frobt.2025.1586033

Modeling arbitrarily applicable relational responding with the non-axiomatic reasoning system: a Machine Psychology approach

Robert Johansson*

Department of Psychology, Stockholm University, Stockholm, Sweden

Arbitrarily Applicable Relational Responding (AARR) is a cornerstone of human language and reasoning, referring to the learned ability to relate symbols in flexible, context-dependent ways. In this paper, we present a novel theoretical approach for modeling AARR within an artificial intelligence framework using the Non-Axiomatic Reasoning System (NARS). NARS is an adaptive reasoning system designed for learning under uncertainty. We introduce a theoretical mechanism called *acquired relations*, enabling NARS to derive symbolic relational knowledge directly from sensorimotor experiences. By integrating principles from Relational Frame Theory—the behavioral psychology account of AARR—with the reasoning mechanisms of NARS, we conceptually demonstrate how key properties of AARR (mutual entailment, combinatorial entailment, and transformation of stimulus functions) can emerge from NARS's inference rules and memory structures. Two theoretical demonstrations illustrate this approach: one modeling stimulus equivalence and transfer of function, and another modeling complex relational networks involving opposition frames. In both cases, the system logically demonstrates the derivation of untrained relations and context-sensitive transformations of stimulus functions, mirroring established human cognitive phenomena. These results suggest that AARR—long considered uniquely human—can be conceptually captured by suitably designed AI systems, emphasizing the value of integrating behavioral science insights into artificial general intelligence (AGI) research. Empirical validation of this theoretical approach remains an essential future direction.

KEYWORDS

artificial general intelligence (AGI), arbitrarily applicable relational responding, operant conditioning, Non-Axiomatic Reasoning System (NARS), machine psychology, adaptive learning

1 Introduction

Human intelligence is marked by an extraordinary capacity for symbolic reasoning—the ability to understand and manipulate symbols (words, ideas, abstract concepts) and their relationships in a flexible manner. An aspect of this flexibility is the capability to derive new relationships between symbols without direct training, purely based on their contextual relations. In cognitive and behavioral psychology, this phenomenon is captured by the concept of Arbitrarily Applicable Relational Responding (AARR), which underlies human language and higher cognition (Hayes et al., 2001; Hayes et al., 2021). AARR refers to the learned behavior of relating stimuli in arbitrary ways (not dictated by the physical properties

of the stimuli, but by contextual cues and social learning). For example, once a child learns that the spoken word "dog" refers to an actual furry pet, the child responds to the word as if it is functionally equivalent to the animal itself—experiencing excitement or happiness when hearing the word, similar to encountering the dog. Such symbolic equivalence is not determined by physical similarity but by relational learning. Derived relational responding of this type is considered a hallmark of human language and reasoning, enabling everything from understanding metaphors to performing complex analogies.

While humans readily perform AARR, instantiating this ability in artificial intelligence (AI) systems remains a formidable challenge. Traditional symbolic AI systems typically rely on explicitly programmed logic rules or axioms, and machine learning systems (like deep neural networks) often require vast amounts of data and struggle with extrapolating knowledge in the absence of direct examples. Achieving human-like symbolic reasoning in a machine calls for an approach that can learn relational patterns from a few examples and generalize them in a context-sensitive way, much as humans do. In other words, we seek an AI that can learn how to relate rather than being pre-programmed with all possible relations.

In this paper, we propose that AARR can be effectively modeled within a particular AI framework known as the Non-Axiomatic Reasoning System (NARS). NARS is an AI reasoning architecture designed to operate under the real-world constraints of insufficient knowledge and resources (i.e., it does not assume a closed, complete set of axioms or unlimited processing power) (Wang, 2013; Wang, 2022). Instead of a fixed logic, NARS uses an adaptive logic (Non-Axiomatic Logic, NAL) that allows it to learn from experience, update its beliefs probabilistically, and make plausible inferences even when knowledge is incomplete. These features make NARS a strong candidate for modeling the emergent, learned relations that characterize AARR.

The key contribution of this work is to demonstrate a computational method for describing human-like symbolic reasoning (AARR) in a machine by utilizing NARS's capabilities. We integrate theoretical insights from Relational Frame Theory (RFT) (Hayes et al., 2001; Hayes et al., 2021) — the behavioral theory that explicates AARR—with the algorithmic machinery of NARS. We propose a novel theoretical mechanism called *acquired relations*, enabling NARS to derive symbolic relational knowledge directly from sensorimotor experiences. In doing so, we show that an AI system can learn and derive relationships among symbols in a manner analogous to human relational learning. This integration provides a framework for studying and implementing cognitive phenomena like language and abstract reasoning in AI. Importantly, our approach goes beyond purely mechanistic or narrow AI methods: rather than training a black-box neural network on vast relational datasets, we employ a functional approach grounded in how relations are learned and used by humans (Johansson, 2024a). This allows the system to capture the contextual control and generalizability of human relational responding.

This integrative approach aligns with the broader interdisciplinary perspective of *Machine Psychology* (Johansson, 2024a; Johansson, 2024b), which systematically applies principles from learning psychology—such as operant conditioning, generalized identity matching, and functional equivalence—to artificial intelligence architectures, aiming to replicate increasingly complex cognitive phenomena in machines (See Table 1 for an overview of how the present research fits with previously conducted studies).

We validate our approach with two experimental paradigms inspired by human studies. The first is a stimulus equivalence task involving three groups of stimuli and tests for derived symmetric and transitive relations, as well as a demonstration of the transformation of stimulus function (e.g., if one stimulus in a set is given a certain meaning or consequence, the others derived to be equivalent to it also reflect that meaning) (Hayes et al., 1987). The second is an oppositional relational network task, where the system learns a network of "opposite" relations (a case of a more complex relational frame) and we examine how this leads to emergent relations and transformations of function consistent with what is observed in human experiments on relational framing of opposites (Roche et al., 2000).

The remainder of this article is organized as follows. Section 2 provides background on Arbitrarily Applicable Relational Responding, the Non-Axiomatic Reasoning System, and our research approach—Machine Psychology. Section 3 reviews related work, contrasting our perspective with other AI and cognitive modeling efforts. Section 4 introduces our theoretical framework, explaining how acquired relations enable modeling of AARR within NARS. Section 5 outlines the methodology behind our illustrative theoretical demonstrations, and Section 6 summarizes their key results, with detailed conceptual derivations provided in the Supplementary Material. Finally, Section 7 discusses broader implications for artificial general intelligence and cognitive science, and outlines directions for future empirical research. Collectively, these contributions establish a theoretical foundation for the empirical study of relational responding in adaptive AI systems.

2 Theoretical background

2.1 Arbitrarily applicable relational responding

Arbitrarily Applicable Relational Responding (AARR) is a concept from behavioral psychology that refers to a general pattern of learned behavior: responding to the relation between stimuli rather than just the stimuli themselves, and doing so in a way that is not determined by the stimuli's physical properties but by contextual cues and history of reinforcement (Hayes et al., 2001; Hayes et al., 2021). This idea is central to Relational Frame Theory (RFT), a modern behavioral theory of language and cognition (Hayes et al., 2001; Hayes et al., 2021). According to RFT, virtually all of human language and higher cognition is founded upon AARR—the ability to treat different stimuli as related along various dimensions (e.g., *same*, *different*, *greater than*, *opposite*, etc.) purely as a result of learned context, not because of any inherent relationship in their physical features.

Three key properties define AARR and distinguish it from simple associative learning.

1. Mutual Entailment: This is the bidirectionality of derived relations. If a person learns a relation in one direction (e.g., A is larger than B), they can derive the relation

TABLE 1 Overview of psychological processes, NARS mechanisms, layers (from Wang, 2013), and references.

Psychological process	NARS mechanisms	NARS layers	References
Operant conditioning	Temporal reasoning and procedural reasoning	7–8	(Johansson, 2024b)
Generalized identity matching	+Abstraction	+6	(Johansson et al., 2023)
Functional equivalence	+Implications	+5	(Johansson et al., 2024)
Arbitrarily applicable relational eesponding	+Acquired relations	+4	This study

in the opposite direction (B is smaller than A) without direct training (Luciano et al., 2007). In classical terms, mutual entailment encompasses symmetric relations (if $A = B$, then $B = A$) and the inverses of asymmetrical relations (if $A > B$, then $B < A$) in a generalized way. Notably, the derived relation might not be identical in form (for instance, *larger than* vs *smaller than* are inverse relations rather than exactly the same), but they are mutually implied by each other given the contextual cues (such as the contextual cue for comparison).

2. Combinatorial Entailment: This is the ability to derive new relations from combinations of learned relations. For example, if one learns that A is related to B, and B is related to C, one can often derive a relation between A and C, depending on the nature of the relation. In the simplest case, if $A = B$ and $B = C$ (coordination relations), then one can derive $A = C$ (equivalence). If $A > B$ and $B > C$ (a comparative relation of "more than"), one can derive $A > C$ ("A is more than C"). These are akin to transitive inferences, but RFT uses the term *combinatorial* entailment to emphasize that the new relation emerges from the combination of two or more other relations.
3. Transformation of Stimulus Function: Perhaps the most distinctive aspect, this refers to the way the functions of stimuli (their meaning, emotional valence, or behavioral effects) can change based on the relations they participate in (Dymond and Rehfeldt, 2000). In other words, if two stimuli are related in a certain way, any psychological function attached to one stimulus (like being pleasant, having a certain name, evoking a specific response) can be transferred to the other stimulus in accordance with their relation. For instance, suppose a person is taught that stimulus A is equivalent to stimulus B ($A = B$, a coordination relation), and separately, stimulus A acquires a particular function (e.g., A is paired with a reward or labeled as "good"). Then, without additional training, the person may treat stimulus B as also having that function (finding B pleasant or "good"), because B is in the same equivalence class as A. If the relation is one of opposition, the functions might transfer in an opposite manner (e.g., if A is opposite to B, and A is associated with "good," B might be seen as "bad") (Roche et al., 2000). Transformation of function demonstrates how relational learning can govern the meaning of symbols in context.

An example can illustrate these principles. Imagine a scenario in a coffee shop: A newcomer is told that "Espresso is stronger than Americano, and Americano is stronger than Caffé au Lait." From just this information, the person can derive that Espresso is stronger than Caffé au Lait, and conversely, Caffé au Lait is weaker than Espresso (combinatorial entailment and mutual entailment for the comparative frame). Now, suppose the person actually tastes an Americano and finds it strong and bitter. That experience may attach a function (strong flavor) to Americano. Due to the relational network, the person might now expect that Espresso (which was said to be stronger than Americano) has an even stronger taste, and that Caffé au Lait (weaker than Americano) has a milder taste, even though they have never tasted Espresso or Caffé au Lait. This is a transformation of stimulus function across a comparative relation network: the direct experience with one item (Americano) transformed the anticipated qualities of the related items (Espresso, Caffé au Lait) in line with the learned relations.

Relational Frame Theory has identified numerous types of relational patterns (called *relational frames*) that humans can learn. Some prominent examples include frames of *coordination* (equivalence/sameness), *distinction* (different from), *comparison* (more than/less than as in the coffee strength example), *opposition*, *hierarchy* (e.g., category membership relations, like "X is a kind of Y"), *temporal* (before/after), *spatial* (here/there), and *deictic* (I/you, now/then, here/there, which involve perspective) (Hayes et al., 2001; 2021). All these frames share the properties of mutual and combinatorial entailment and can lead to transformations of function, though the exact nature of the entailments depends on the frame.

It is important to note that AARR is considered an *operant behavior*, meaning it is learned through a history of reinforcement and context, rather than being an innate or automatic reflex (Hayes et al., 2021). Crucially, according to RFT, derived relational responding (such as mutual entailment, combinatorial entailment, and transformation of function) is established via *multiple exemplar training (MET)*, a well-documented learning process through which individuals are exposed to a variety of relational examples until relational responding generalizes to new, untrained examples without direct reinforcement (Luciano et al., 2007; Hayes et al., 2021). Thus, explicitly training relational patterns initially is fully consistent with RFT, and subsequent relational responding is considered "emergent" precisely because it generalizes beyond reinforced examples due to this learning history. The term "arbitrarily applicable" emphasizes that any stimuli, regardless of their formal properties, can be related in any way, given the appropriate training context. Humans, especially those with language ability, seem uniquely capable of this kind of learning (Devany et al., 1986). Indeed, research has shown that stimulus equivalence (a basic form of AARR focusing on sameness) reliably

appears in humans but not in most non-human animals without language training, with only rare exceptions (Schusterman and Kastak, 1993). This link between language and AARR suggests that a capacity for relational responding is a defining feature of higher cognition.

Relational Frame Theory provides a perspective on general intelligence as well. Rather than viewing intelligence as a monolithic IQ or a fixed set of problem-solving abilities, RFT suggests intelligence involves a rich repertoire of relational skills (Cassidy et al., 2016; Hayes et al., 2021). From this viewpoint, improving one's ability to learn and manipulate complex relational networks should enhance cognitive performance. Studies have found that training individuals on relational tasks can increase scores on standard intelligence tests (Cassidy et al., 2016). Programs like *SMART* (Strengthening Mental Abilities with Relational Training) and *PEAK* (Promoting the Emergence of Advanced Knowledge) aim to boost cognitive and language abilities by systematically exercising relational responding abilities (Dixon et al., 2017).

In summary, AARR, as characterized by RFT, captures the flexibility, generativity, and context-sensitivity of human symbolic reasoning. Modeling this phenomenon in an AI system requires that the system can represent relations between symbols, infer new relations from old, and dynamically update what symbols mean based on relational context. Next, we discuss NARS, which we propose as a suitable candidate for this challenge.

2.2 Non-Axiomatic Reasoning System (NARS)

The Non-Axiomatic Reasoning System (NARS) is an AI system and cognitive architecture developed by Pei Wang (Wang, 2013; Wang, 2022) with the goal of realizing a form of general intelligence that operates effectively under real-world constraints. The name "non-axiomatic" reflects that NARS does not assume a predefined, complete set of axioms or truths about the world; instead, it must learn and reason non-axiomatically, meaning all its knowledge is gleaned from experience and is always revisable. NARS was built on the recognition that an intelligent agent in the real world must cope with insufficient knowledge and insufficient resources (a principle Wang abbreviates as AIKR: Assumption of Insufficient Knowledge and Resources (Wang, 2019)). Unlike classical logic systems that are brittle outside of their given axioms, NARS is adaptive and is constantly updating its beliefs and strategies as new information comes in, somewhat akin to a human continually learning and adjusting their understanding.

At the core of NARS is an AI reasoning framework called Non-Axiomatic Logic (NAL). NAL is a formal logic that extends term logic (a kind of logic dealing with relationships between terms or concepts) and is probabilistic in nature. NARS uses an internal language, Narsese, to represent knowledge. All pieces of knowledge in NARS are expressed as statements in Narsese, which typically have a subject and a predicate and a copula connecting them (the copula defines the type of relation between subject and predicate). The simplest form is an inheritance relation "$S \rightarrow P$" meaning "S is a kind of P" or "S implies P" in a category sense. For example, one could represent "Tweety is a bird" as `Tweety` → `Bird`, and "Birds are animals" as `Bird` → `Animal`. NAL can then derive `Tweety` → `Animal` by inference (a kind of syllogism) (Wang, 2013). In addition to inheritance, Narsese includes other basic copulas such as similarity (noted as ↔ in Narsese, meaning two terms are similar or equivalent in some sense), implication (→ with different context indicating temporal or causal implication), and equivalence (⇔ for bi-conditional statements). Through combinations of these, NARS can represent a wide variety of knowledge, including rules like "if X happens then Y tends to happen" (an implication), or "Concept A is similar to Concept B" (a similarity statement).

Crucially, every statement in NARS carries a measure of uncertainty. NARS does not use binary true/false assignments; instead, each piece of knowledge has a truth value with two parameters: *frequency* (a measure akin to probability based on how often the relation has been true in experience) and *confidence* (reflecting the amount of evidence available) (Hammer, 2022; Wang, 2022). This allows NARS to reason under uncertainty and update its beliefs as new evidence arrives. For example, if initially NARS has little evidence about "Tweety can fly," it might assign it a low confidence. If many observations confirm it, the confidence (and perhaps the frequency) increases. See the Supplementary Material for more information regarding frequency and confidence.

Another distinguishing feature of NARS is its approach to resource constraints. NARS operates in real-time and has a limited "budget" for attention and memory. It cannot consider all knowledge all the time. Instead, it uses a priority mechanism to decide which tasks (questions, goals, new knowledge) to process next, based on factors like urgency and relevance. This ensures that at any given moment, the system focuses on the most pertinent information, allowing it to scale to larger problems by not getting bogged down in less relevant details.

Recent implementations of NARS include OpenNARS and specifically a variant called OpenNARS for Applications (ONA) (Hammer and Lofthouse, 2020). ONA is tailored for integration into practical applications, including robotics. It extends the basic NARS framework with sensorimotor capabilities, meaning it can handle input from sensors and send output to actuators (motors) as part of its reasoning. This is done by treating sensorimotor events also as terms in the language (for instance, a sensory observation or a motor command can be a term that participates in statements). In ONA, the reasoning engine is capable of doing temporal inference, understanding sequences of events and causality. Temporal relations in Narsese might be represented with additional notation - for example, $A \Rightarrow B$ might denote events A and B happening in sequence. ONA's design includes components like event buffers, concept memory, and distinct inference processes for different types of tasks (e.g., some for immediate reactions, some for long-term learning) (Hammer and Lofthouse, 2020; Hammer, 2022).

For the purposes of this work, what is important is that NARS (and ONA) provides.

- A flexible knowledge representation that can express arbitrary relations between symbols (via terms and copulas in Narsese).
- Inference rules that can derive new relationships from known ones, analogous to the entailments described in RFT. For example, NARS can perform syllogistic inference (if $A \rightarrow B$ and $B \rightarrow C$, derive $A \rightarrow C$) and inductive inference (generalizing

or specializing relations based on evidence), which parallel combinatorial entailment in AARR.
- The ability to incorporate new knowledge on the fly and revise existing knowledge, which is essential for any learning system attempting to acquire relational behavior through training.
- The ability to handle context and switch between tasks, somewhat akin to how contextual cues in AARR determine which relation applies. In NARS, context is handled through its concept activations and the specific questions posed to the system; it is not identical to the notion of contextual cues in RFT, but NARS can take context into account by treating it as just another piece of information in the premise of a statement or rule.

In short, NARS can be seen as a unified cognitive model that does not separate reasoning, learning, memory, and perception into different modules; the same underlying logic and control mechanism handles all these functions (Wang et al., 2022). This makes it very appealing for modeling complex cognitive phenomena like AARR, because we do not need to bolt together separate systems for learning relations and for reasoning about them—NARS does both in one framework. The challenge is to design the right way to present relational training to NARS and possibly to extend NARS with any additional mechanisms so that it can exhibit mutual and combinatorial entailment and transformation of functions in a manner comparable to humans.

2.3 Machine psychology: bridging learning psychology and adaptive AI

Machine Psychology is an interdisciplinary framework that integrates learning psychology with adaptive AI systems, such as NARS, to explore the emergence of cognitive behaviors in artificial agents (Johansson, 2024a; Johansson, 2024b). This approach systematically investigates increasingly complex learning processes, drawing from operant conditioning, generalized identity matching, and functional equivalence, which are fundamental to relational cognition. In Table 1, we clarify how this systematic approach has been carried out in previous studies.

In this work, we assume that the system is interacting with the environment using different sensors. A key sensor that will be used throughout the entire paper is the assumption of a location sensor. Objects perceived by the vision system would using this model all be assigned a location. The labels `sample`, `left`, `right`, etc., are totally arbitrary. They are chosen by the designer and are only labels used to indicate that different objects are perceived at different locations.

We could also imagine that the system is equipped with a color sensor, and is interacting with a Matching-to-sample procedure. For example, as illustrated in Figure 1, something red is in the sample position, something green is to the left, and something blue to the right. This could be described that the only "eyes" that the system have are location and color, meaning that other object properties like shape and size couldn't be perceived by that system.

The way we represent such interactions with the world in this paper is like the following:

```
<(sample * red) --> (loc * color)>. :|:
```

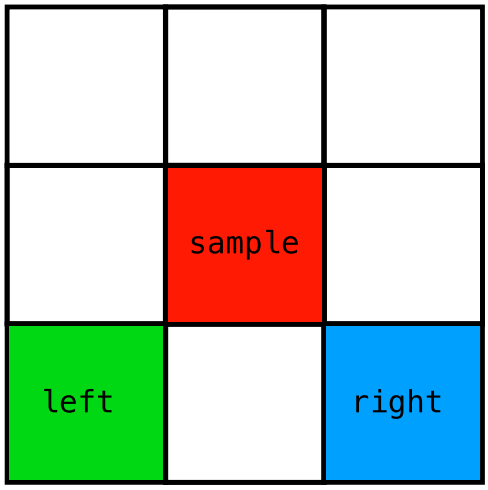


FIGURE 1
An example scene where the system perceives three different colors at three different locations.

```
<(left * green) --> (loc * color)>. :|:
<(right * blue) --> (loc * color)>. :|:
```

The scene is described by two temporal statements (as indicated by `:|:`). Perceiving a green object to the left can be described as an interaction between perceiving to the left, and perceiving green. Hence, the statement `<(left * green) --> (loc * color)>` can be seen as a composition of `<left --> loc>` and `<green --> color>`. This encoding of object properties at certain locations will be used throughout this paper. Importantly, also an OCR detector will be assumed in the experiments carried out in the present study, leading to interactions as the one illustrated below.

```
<(sample * A1) --> (loc * ocr)>. :|:
<(left * B1) --> (loc * ocr)>. :|:
<(right * B2) --> (loc * ocr)>. :|:
```

For details regarding the experimental setup used in the research described below, see the section in the Supplementary Material, that clarifies Narsese syntax and key concepts.

2.3.1 Operant conditioning with NARS

The foundation of Machine Psychology is built on operant conditioning, a fundamental demonstration of adaptive behavior (Johansson, 2024b). In our research, NARS was exposed to operant contingencies where behaviors were reinforced based on temporal and procedural reasoning. This enabled NARS to learn through interaction with its environment, adjusting actions based on feedback, similar to how organisms learn in response to consequences. The results demonstrated that NARS could acquire and refine behaviors through reinforcement, providing an essential basis for more advanced relational learning.

```
<(left * blue) --> (loc * color)>. :|:
<(right * green) --> (loc * color)>. :|:
G! :|: // Establish G as a goal
// Executed with motor babbling:
// ^select executed with args
({SELF} * right)
```

```
G. :|: // Provide G as a consequence
// Derived with frequency 1, and confidence 0.19:
// <(<(right * green) --> (loc * color)> &/ <({SELF} * right) --> ^select>) =/> G>.
```

2.3.2 Generalized identity matching with NARS

Building upon operant conditioning, our research extended into generalized identity matching, which involves recognizing and responding to identity relations across varying stimuli (Johansson et al., 2023). This required NARS to utilize complex learning mechanisms, including abstraction and relational generalization. By introducing an abstraction mechanism to NARS, we enabled it to derive identity relations beyond explicit training examples, mirroring human cognitive abilities in symbolic matching tasks. The results showed that NARS could generalize identity relations to novel stimuli, demonstrating an emergent form of relational reasoning.

Let's say that the system was exposed to the following NARS statements in the training phase.

```
<(sample * blue) --> (loc * color)>. :|:
<(left * green) --> (loc * color)>. :|:
<(right * blue) --> (loc * color)>. :|:
G! :|:
```

NARS could execute `match` with `sample` and `right` (from motor babbling or a decision based on previous experience), which would be considered correct, and hence the feedback `G. :- :` would be given to NARS, followed by 100 time steps. Only from this single interaction, NARS would form both a specific and a general hypothesis.

```
<((<(sample * blue) --> (loc * color)> &/
<(right * blue) --> (loc * color)>) &/
<({SELF} * (sample * right)) --> ^match>) =/> G>
// frequency: 1.00, confidence: 0.15
<((<(#1 * #2) --> (loc * color)> &/
<(#3 * #2) --> (loc * color)>) &/
<({SELF} * (#1 * #3)) --> ^match>) =/> G>
// frequency: 1.00, confidence: 0.15
```

2.3.3 Functional equivalence with NARS

Further advancing Machine Psychology, we explored functional equivalence, a process in which stimuli become interchangeable in guiding behavior due to shared functional properties (Johansson et al., 2024). This study introduced additional inference mechanisms into NARS, allowing it to derive new relations based on implications and acquired equivalences. Functional equivalence is critical for understanding how abstract categories are formed and used in problem-solving. Our findings indicate that NARS can establish and apply functional equivalence relations, effectively transferring learned functions between distinct but related stimuli.

```
<(s1 * A1) --> (loc * ocr)>. :|:
G! :|:
// Executed with motor babbling
<({SELF} * R1) --> ^press>. :|:
G. :|:
// Derived
<(<(s1 * A1) --> (loc * ocr)> &/
<({SELF} * R1) --> ^press>) =/> G>.
100
<(s1 * A2) --> (loc * ocr)>. :|:
G! :|:
// Executed same operation with motor babbling
<({SELF} * R1) --> ^press>. :|:
G. :|:
// Derived
<(<(s1 * A2) --> (loc * ocr)> &/
<({SELF} * R1) --> ^press>) =/> G>.
```

Since the system derived two contingencies that only differed in the pre-condition, statements like the following (functional equivalence) would also be derived.

```
<<($1 * A1) --> (loc * ocr)> ==> <($1 * B1) --> (loc * ocr)>>.
<<($1 * B1) --> (loc * ocr)> ==> <($1 * A1) --> (loc * ocr)>>.
```

These studies collectively illustrate the progression from simple operant conditioning to complex relational cognition, reinforcing Machine Psychology as a viable framework for advancing artificial general intelligence (AGI). An overview of the systematic approach Machine Psychology has taken, can be seen in Table 1. By systematically integrating behavioral learning principles with adaptive AI reasoning, this approach contributes to the development of more flexible, human-like intelligence in machines.

3 Related work

Integrating principles of human cognition and learning into AI systems is a growing interdisciplinary endeavor. However, Relational Frame Theory (RFT) and its core concept of Arbitrarily Applicable Relational Responding (AARR) have seen relatively little application in mainstream AI research. Most approaches to relational reasoning in AI have taken alternative paths.

3.1 Symbolic AI and knowledge graphs

Traditional symbolic reasoning systems, such as knowledge graph inference engines and logic-based AI, typically represent relations axiomatically (Lenat, 1995; Rosenbloom et al., 2016). These systems utilize explicitly predefined relational structures (e.g., ontological relationships like "isFatherOf" being inverse to "isChildOf"). They do not usually learn these relations dynamically but rely instead on manually crafted knowledge. In contrast, the proposed NARS-based approach aims at learning arbitrary relations from experience, enabling dynamic derivation of novel relations without predefined axioms.

3.2 Machine learning for relational tasks

In the machine learning domain, methods such as relational reinforcement learning, graph neural networks, and transformer-based models excel at extracting patterns from relational datasets.

For example, DeepMind's Relation Networks can effectively learn relational structures to answer visual-spatial questions from large-scale training data (Santoro et al., 2017). However, these data-driven methods typically require substantial training examples and may not guarantee key relational properties such as mutual or combinatorial entailment. Furthermore, these methods often lack interpretability and struggle with few-shot generalization—a core strength of human cognition that NARS aims to model by deriving relational structures adaptively from minimal and context-sensitive experiences.

3.3 Bayesian approaches to relational learning

Bayesian methods, including probabilistic programming and Bayesian relational modeling, represent relational structures while also modeling uncertainty (Nitti et al., 2016; Tenenbaum et al., 2011). These approaches are highly effective in generalizing from limited data, but they typically depend on predefined model structures and well-defined priors. As a result, dynamically deriving novel relational structures purely from interaction or flexibly adapting to context-sensitive relations can be challenging. By contrast, our NARS-based framework inherently constructs relational structures directly from interaction and accommodates dynamic, context-dependent inference without reliance on extensive predefined priors.

3.4 Statistical relational learning and neurosymbolic AI

Recent advances in Statistical Relational Learning (SRL) and neurosymbolic AI methods integrate symbolic logic with statistical and neural learning techniques (Marra et al., 2024). These hybrid methods effectively handle relational inference tasks by leveraging symbolic representation and data-driven learning. However, SRL methods typically require large datasets and predefined structures, potentially limiting their adaptability in low-data or dynamically evolving contexts. Our approach utilizing NARS offers a complementary perspective by emphasizing adaptive reasoning and minimal-data learning, targeting scenarios that demand rapid relational inference from limited interactions.

3.5 Inductive logic programming

Inductive Logic Programming (ILP) is another well-established paradigm for symbolic relational learning, focusing on deriving relational rules from structured data (Cropper and Dumančić, 2022). Recent ILP applications have successfully modeled cognitive processes in robotic systems, enabling robots to generalize relational tasks from expert feedback (Meli and Fiorini, 2025). While powerful, ILP generally relies on explicitly defined logical frameworks and structured training examples. In contrast, our proposed integration of NARS and RFT uniquely emphasizes adaptive, context-sensitive relational learning, minimizing reliance on predefined logic templates or extensive datasets.

3.6 Computational approaches inspired by RFT

Few computational approaches explicitly model AARR as defined by RFT. Early computational models attempted to simulate stimulus equivalence and relational responding through neural network approaches (Barnes and Hampson, 1993; Cullinan et al., 1994). These connectionist methods successfully modeled basic relational properties such as symmetry and transitivity but typically required extensive training data and had limited scalability to complex relational frameworks. Although computational modeling of stimulus equivalence remains active (Tovar et al., 2023), modeling of broader AARR principles beyond stimulus equivalence is rare, with notable exceptions including recent works by Edwards et al. (2022); Edwards (2024).

In summary, relational reasoning remains a vibrant area within AI research, yet the challenge of dynamically learning arbitrary, contextually flexible relational structures with minimal training data remains largely unmet. Our proposed NARS-based framework directly addresses this gap. To the best of our knowledge, this study is the first to conceptually demonstrate how mutual entailment, combinatorial entailment, and transformation of functions—key properties of AARR—can emerge within a unified symbolic reasoning system. This theoretical foundation sets the stage for future empirical validations and positions NARS as a promising candidate for adaptive, human-like relational reasoning.

4 Theoretical framework: modeling AARR with NARS

To enable the modeling of Arbitrarily Applicable Relational Responding (AARR) within OpenNARS for Applications (ONA), we introduce a novel mechanism called *acquired relations*. Currently, ONA's reasoning is based primarily on sensorimotor contingencies; however, according to NARS theory (NAL Definition 8.1 in Wang (2013)), relational terms (*products*) can equivalently be represented as compound terms of inheritance statements. This theoretical notion has not yet been implemented in ONA, and its introduction would allow the system to derive relational statements directly from learned sensorimotor contingencies.

Within NARS theory, a learned contingency such as.

```
<((<A1 --> p1> &/ <B1 --> q1>) &/ ^left)
=/> G>.
```

can yield an *acquired relation*, formally represented as.

```
<(A1 * B1) --> (p1 * q1)>.
```

In the notation employed here, learned sensorimotor contingencies often take the form.

```
<(sample * red) --> (loc * color)> &/
<(left * blue) --> (loc * color)> &/
<({SELF}    * (sample * left)) --> ^match>
=/> G>.
```

Following our approach, this yields two distinct relational terms—one describing the relation between stimulus properties (colors), and another describing the relational structure of stimulus locations.

```
<(red * blue) --> (color * color)> &&
<(sample * left) --> (loc * loc)>
```

To avoid a *combinatorial explosion*, i.e., an exponential growth in derived terms and inferences, the introduction of acquired relations is carefully restricted. Specifically, new relations are generated only when procedural operations within contingencies are actively executed by the system. This targeted triggering ensures computational efficiency while maintaining functional generality.

Acquired relations can be combined with *implications*, another core element in NARS theory (see statement-level inference in Wang (2013)), allowing for generalized, context-sensitive reasoning. For example, from the acquired relations shown previously, the following implications can be derived.

```
<(red * blue) --> (color * color)> &&
<(sample * left) --> (loc * loc)> ==>
<(sample * red) --> (loc * color)> &/
<(left * blue) --> (loc * color)> &/
<({SELF} * (sample * left)) --> ^match> =/> G>.
```

More generally, implications abstracted with variables take this form.

```
<($1 * $2) --> (color * color)> &&
<($3 * $4) --> (loc * loc)> ==>
<($3 * $1) --> (loc * color)> &/
<($4 * $2) --> (loc * color)> &/
<({SELF} * ($3 * $4)) --> ^match> =/> G>.
```

This framework can be understood as a grounding mechanism whereby abstract relations (e.g., color-color) become anchored in concrete sensorimotor experiences. This allows NARS to dynamically transition from basic, animal-like contingency learning towards symbolic, human-like reasoning capabilities.

When multiple abstract relational templates or rules could apply during inference, NARS selects among these templates by prioritizing the rule with the highest *truth expectation* (Hammer and Lofthouse, 2020). Truth expectation in NARS is calculated as a function of frequency and confidence associated with previously derived relational implications:

$$exp(f,c) = c \times \left(f - \frac{1}{2}\right) + \frac{1}{2}$$

where frequency (f) represents the proportion of positive evidence relative to the total evidence, and confidence (c) reflects the degree of evidential support based on the total amount of evidence (Hammer and Lofthouse, 2020). Thus, inference proceeds using the relational rule with the strongest combined evidential support, reflecting the system's accumulated relational learning experiences.

Conversely, symbolic-level relational statements can also guide sensorimotor behavior. If a relation such as (*blue* → *yellow*) is symbolically derived, it can then inform decision-making in novel situations via the implications described above, provided relevant locational relations (e.g., (*sample* → *right*)) are established through direct interaction with the environment.

The concept of acquired relations is general and not restricted to matching-to-sample procedures. For example, functional equivalences acquired through interactions with different procedures also lead to relational derivations. Consider the following example.

```
<(<(left * green) --> (loc * color)> &/
<({SELF} * left) --> ^select>) =/> G>
100
<(<(left * blue) --> (loc * color)> &/
<({SELF} * left) --> ^select>) =/> G>
// Derived functional equivalence:
<(left * green) --> (loc * color)> <=>
<(left * blue) --> (loc * color)>
```

This equivalence, in turn, can support acquired relational implications.

```
<(green * blue) --> (color * color)> &&
<(left * left) --> (loc * loc)> ==>
<(left * green) --> (loc * color)> <=>
<(left * blue) --> (loc * color)>
// Abstracted form:
<($1 * $2) --> (color * color)> &&
<($3 * $3) --> (loc * loc)> ==>
<($3 * $1) --> (loc * color)> <=>
<($3 * $2) --> (loc * color)>
```

This flexibility aligns closely with contemporary learning psychology perspectives, which argue that any regularity—such as stimulus pairing or common roles within contingencies—can serve as a contextual cue for relational responding (De Houwer and Hughes, 2020; Hughes et al., 2016).

In the following section, we detail specific experimental paradigms designed to validate and explore the capabilities enabled by these modeling extensions.

5 Illustrative theoretical demonstrations

The following sections present conceptual scenarios illustrating logical derivations rather than empirical experiments. These demonstrations serve as theoretical proofs-of-concept, designed to illustrate how the proposed NARS extensions could enable Arbitrarily Applicable Relational Responding (AARR). Quantitative performance metrics (e.g., accuracy, F1-score) are not applicable in this purely theoretical context but remain important targets for future empirical evaluations.

Crucially, during all theoretical testing phases reported here, we presented only the goal-event (`G! :|:`) to trigger system choices. We never provided feedback or reinforcement (`G. :|:`) during these tests. Thus, our testing phases strictly followed standard Matching-to-Sample (MTS) procedures used in human relational research, ensuring genuine tests of generalization in the absence of feedback. Please see the Supplementary Material for details.

In alignment with standard Matching-to-Sample procedures used in the human studies we replicate, the spatial positions (left/right) of comparison stimuli were systematically varied and balanced across trials within each training and testing block. This procedure, which has also been employed consistently in our previous experimental research with NARS-based systems (Johansson et al., 2023; Johansson, 2024b), ensures that relational responding could not rely on positional cues.

During training phases, we propose providing feedback in the form of positive reinforcement (`G. :|:`) for correct responses and negative feedback (`G. :|: {0.0 0.9}`) for incorrect responses. In this conceptual framework, negative feedback would reduce the truth expectation of corresponding implications, theoretically decreasing the probability that NARS would repeat incorrect behavior. This approach allows NARS, at a theoretical level, to

adapt relational knowledge based on experience. However, empirical testing of this mechanism remains an essential direction for future research.

We adapted two paradigms from Relational Frame Theory (RFT) literature: the Stimulus Equivalence and Function Transfer task (Task 1; Figure 2) and the Opposition and Function Transformation task (Task 2; Figure 3) (Hayes et al., 1987; Roche et al., 2000). These tasks were modified to conceptually fit the capabilities of NARS. Importantly, these setups were not implemented empirically in OpenNARS for Applications (ONA) (Hammer and Lofthouse, 2020); rather, they are presented here as symbolic analyses intended to illustrate how NARS, when theoretically extended, could account for these forms of relational reasoning.

5.1 Task 1: stimulus equivalence and transfer of function

The design for Task 1 was inspired by the methodology introduced by Hayes et al. (1987). In their original human study, participants underwent four phases: (1) training conditional discriminations, (2) testing for derived equivalence classes, (3) training discriminative stimulus functions on selected class members, and (4) testing whether discriminative functions transferred to other members of the same equivalence classes. Importantly, the original study did not account for participants' prior relational learning history.

In the present study, we included pretraining to establish basic relational skills prior to the main experiments. The study consisted of four phases conducted sequentially.

1. Pretraining of relational networks: This phase explicitly trained foundational relations such as symmetry ($X1 \rightarrow Y1$ and $Y1 \rightarrow X1$), and transitivity ($X1 \rightarrow Y1$, $Y1 \rightarrow Z1$, thus deriving $X1 \rightarrow Z1$).
2. Training conditional discriminations: Using a Matching-to-sample (MTS) procedure, conditional discriminations were trained within two separate stimulus networks: one comprising stimuli $A1$, $B1$, and $C1$, and another comprising $A2$, $B2$, and $C2$.
3. Function training: NARS was trained to execute two discriminative responses: `^clap` when $B1$ was presented as a sample stimulus, and `^wave` when $B2$ appeared as the sample.
4. Testing derived relations and transfer: In the final phase, derived relations within each *ABC* network were tested without feedback, specifically examining whether previously trained discriminative functions (`^clap`, `^wave`) transferred to equivalent stimuli ($C1$, $C2$).

5.2 Task 2: opposition and transformation of function

Task 2 was inspired by the relational methodology of Roche et al. (2000). Roche and colleagues examined how derived relational responses and stimulus functions transformed contextually using "Same" and "Opposite" relational frames. Their human

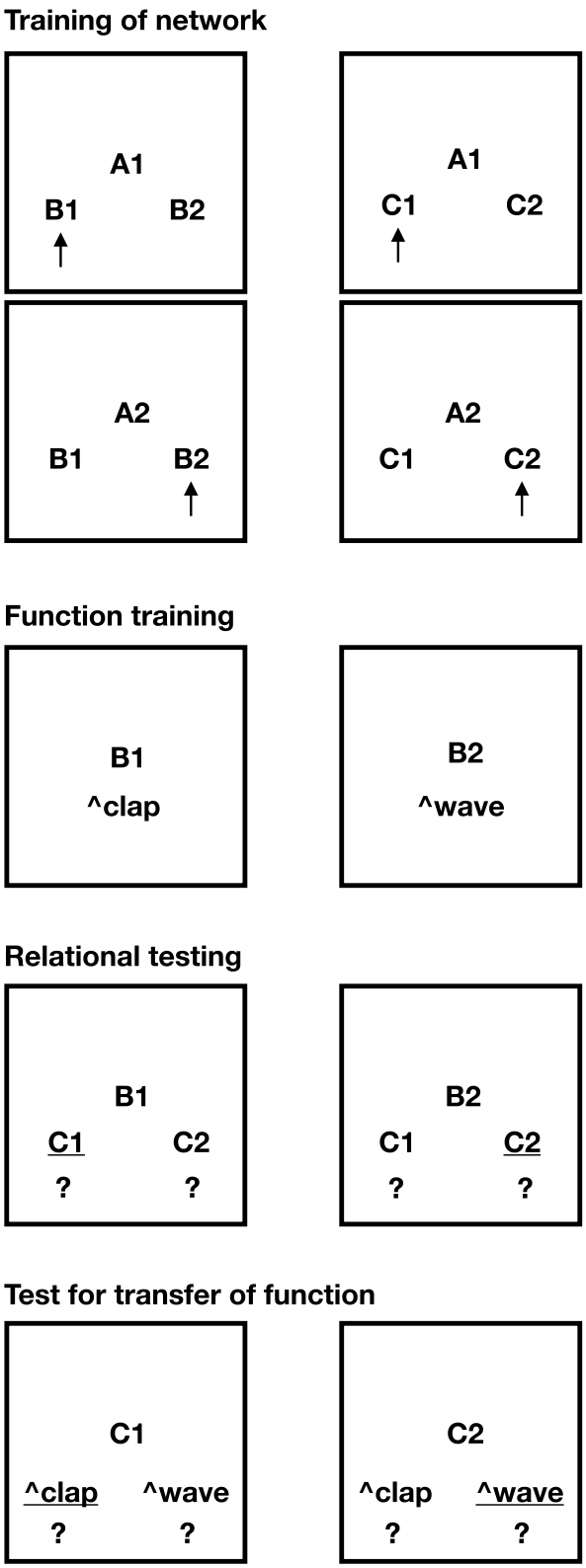


FIGURE 2
Task 1 of this paper. Stimulus equivalence and the transfer of function. The necessary pre-training (Phase 1) is excluded from the picture. Picture shows Phases 2–5 of the task. Underlined options indicate correct choices.

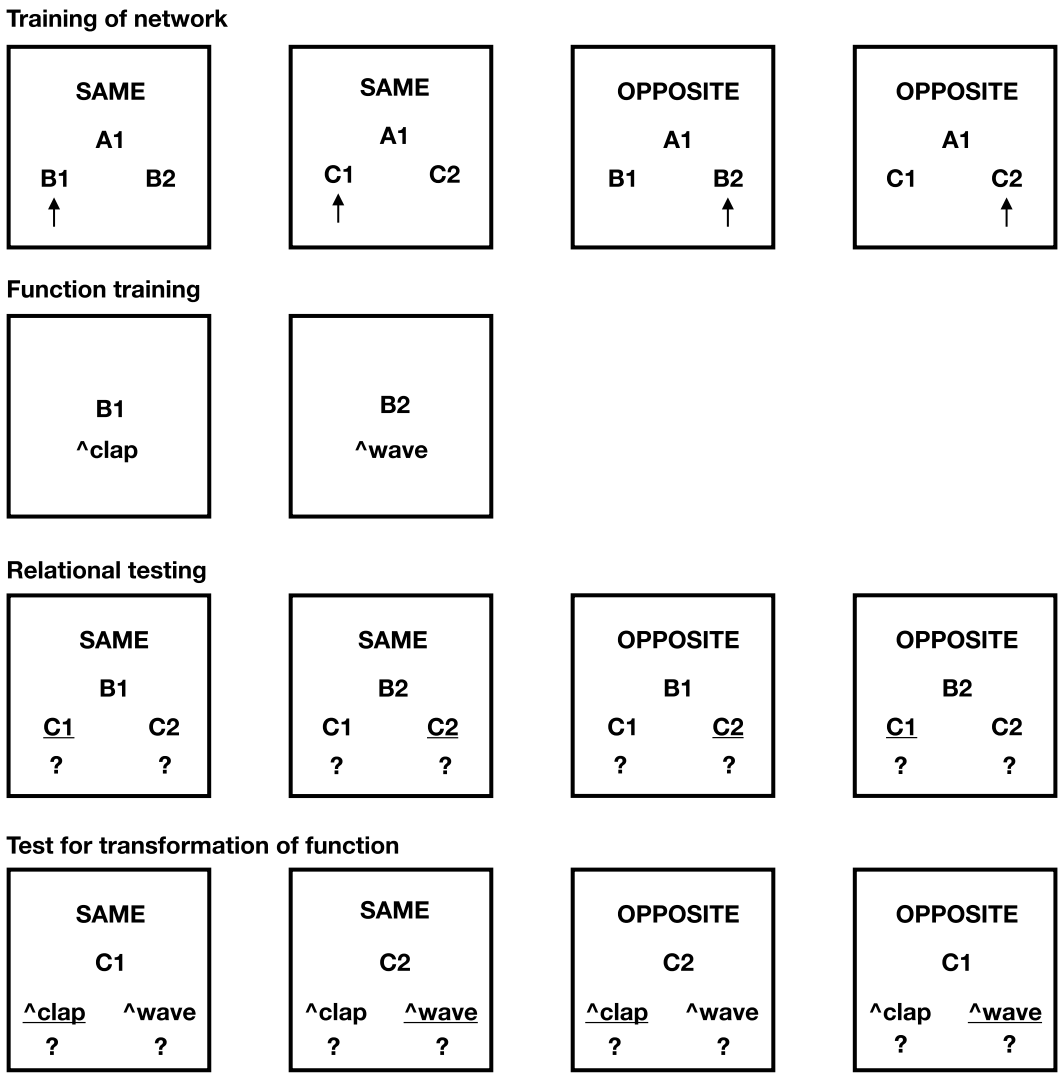


FIGURE 3
Task 2 of this paper. AARR in accordance with opposition and the transformation of function. The necessary pre-training (Phase 1) is excluded from the picture. Picture shows Phases 2–5 of the task. Underlined options indicate correct choices.

participants initially learned operant associations between arbitrary stimuli and actions (e.g., waving, clapping), followed by relational pretraining to establish "Same" and "Opposite" frames. Through training and contextual cueing, participants showed contextually controlled derived responding (e.g., relationally responding "Same" or "Opposite" for specific stimuli) and function transformation.

In the current study, we again included explicit pretraining phases to equip NARS with necessary relational skills. The experimental design comprised five phases.

1. Pretraining of relational frames: This phase explicitly trained "SAME" and "OPPOSITE" relations, establishing mutual entailment (e.g., SAME $X1 \leftrightarrow Y1$, OPPOSITE $X1 \leftrightarrow Y2$) and combinatorial entailment (e.g., SAME $X1 \rightarrow Y1$, SAME $Y1 \rightarrow Z1$, thus deriving SAME $X1 \rightarrow Z1$). Functional equivalence and transfers between symmetry and functional equivalence were also established.
2. Training relational networks: Using the Matching-to-sample (MTS) procedure, relational networks were trained, forming SAME (e.g., $A1 \rightarrow B1$, $A1 \rightarrow C1$) and OPPOSITE ($A1 \rightarrow B2$, $A1 \rightarrow C2$) relations. A second analogous network ($A2$-$B2$-$C2$) was similarly trained.
3. Function training: The system was trained to produce discriminative responses `^clap` (for $B1$) and `^wave` (for $B2$).
4. Testing derived relations and function transformations: In the final phase, derived relations within the SAME/OPPOSITE networks were tested without feedback, specifically examining whether trained functions transformed appropriately across relational contexts. Stimuli tested included combinations such as SAME/$C1$, SAME/$C2$, OPPOSITE/$C1$, and OPPOSITE/$C2$.

6 Theoretical results and conceptual derivations

Given the detailed and extensive nature of the logical derivations underlying these theoretical demonstrations, the full derivations, explicit representations, and step-by-step processes are presented in the Supplementary Material. Here, we summarize the key outcomes of our theoretical demonstrations evaluating whether NARS, with the proposed extensions, can model Arbitrarily Applicable Relational Responding (AARR). The main text thus maintains readability by focusing on the key relational properties (mutual entailment, combinatorial entailment, and transformation of function) that conceptually emerge within the NARS framework.

6.1 Stimulus equivalence and transfer of function

In the first experiment (illustrated in Figure 2), we explored whether NARS logic could model the formation of stimulus equivalence classes and demonstrate the transfer of stimulus functions across related stimuli. Briefly, NARS was theoretically exposed to matching-to-sample (MTS) procedures where conditional relations ($A \rightarrow B$ and $B \rightarrow C$) were trained. Additionally, discriminative functions were assigned to specific stimuli within these relational networks (e.g., stimulus B_1 triggering a `^clap` response, and B_2 a `^wave` response).

Key results included.

- Mutual entailment: NARS successfully derived bidirectional relations (e.g., if trained $A \rightarrow B$, it inferred $B \rightarrow A$).
- Combinatorial entailment: The system correctly inferred indirect relations from explicitly trained ones (e.g., from $A \rightarrow B$ and $B \rightarrow C$, it inferred $A \rightarrow C$).
- Transformation of function: Critically, discriminative functions (e.g., `^clap` and `^wave`) initially trained on B-stimuli were transferred without additional training to C-stimuli through derived equivalence relations, demonstrating a successful relational transfer of stimulus functions.

Thus, NARS logic adequately models essential aspects of stimulus equivalence and function transfer, foundational within Relational Frame Theory (Figure 4; detailed derivations in Supplementary Material Section S1).

6.2 Opposition and transformation of function

In the second experiment (illustrated in Figure 3), we assessed whether NARS logic could model relational networks involving oppositional frames ("SAME" and "OPPOSITE") and the contextual transformation of stimulus functions. Similar to the first task, MTS training was theoretically applied, but now relations involved both SAME and OPPOSITE contexts. After training, discriminative functions were again assigned to specific stimuli within these networks.

Key outcomes included.

- Context-sensitive mutual entailment and combinatorial entailment: NARS derived relations consistent with trained SAME and OPPOSITE relational frames, correctly generalizing from trained examples.
- Transformation of function across oppositional relations: Trained discriminative functions (e.g., `^clap` associated with stimulus B_1, and `^wave` with B_2) were accurately transferred to related stimuli (C_1 and C_2), including appropriate reversal in functions when oppositional relational contexts were applied (e.g., if stimulus pairs were related as OPPOSITE, stimulus functions reversed accordingly).

These results illustrate that NARS logic effectively models complex, contextually controlled transformations of function, consistent with Relational Frame Theory (Figure 5; detailed derivations in Supplementary Material Section S2).

In summary, these theoretical demonstrations confirm that the extended NARS logic is sufficiently powerful and flexible to capture core relational learning phenomena—mutual entailment, combinatorial entailment, and transformation of function—essential for modeling human-like symbolic reasoning and cognition.

7 Discussion

This study demonstrated that the Non-Axiomatic Reasoning System (NARS), extended with mechanisms inspired by Relational Frame Theory (RFT), can successfully model Arbitrarily Applicable Relational Responding (AARR), a cornerstone of human cognition. Through theoretical analysis and logical derivations, we showed how NARS's adaptive logic can capture essential relational learning phenomena without pre-defined axioms or extensive data-driven training. This integration provides a computational framework aligning cognitive science principles with artificial intelligence (AI), underscoring the interdisciplinary potential of Machine Psychology (Johansson, 2024a; Johansson, 2024b) in developing flexible, context-sensitive reasoning systems.

7.1 Summary of theoretical insights

We have shown theoretically that NARS can replicate critical aspects of human-like relational reasoning by modeling Arbitrarily Applicable Relational Responding. Specifically, we demonstrated that.

- NARS exhibits *mutual entailment*, accurately deriving bidirectional relations from trained unidirectional associations.
- It demonstrates robust *combinatorial entailment*, integrating multiple trained relations to correctly infer novel relations.
- It successfully replicates *transformation of stimulus function*, whereby functions (such as specific responses like "clap" or "wave") trained to one stimulus are systematically transferred to other related stimuli without additional direct training.

These findings illustrate that the cognitive mechanisms underlying AARR—once considered unique to biologically evolved cognition—can be conceptually instantiated within a symbolic reasoning system. NARS's capability to learn from minimal,

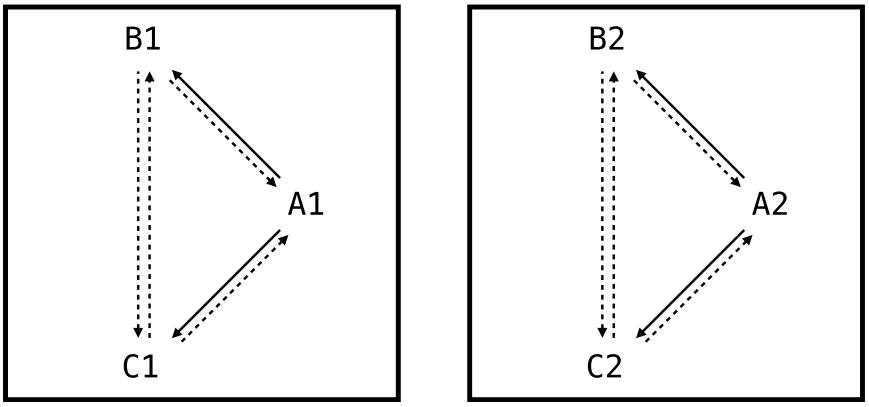


FIGURE 4
The two networks trained as part of the first experiment of this paper. Solid arrows represent relations that are explicitly trained. Dashed arrows represent derived relations.

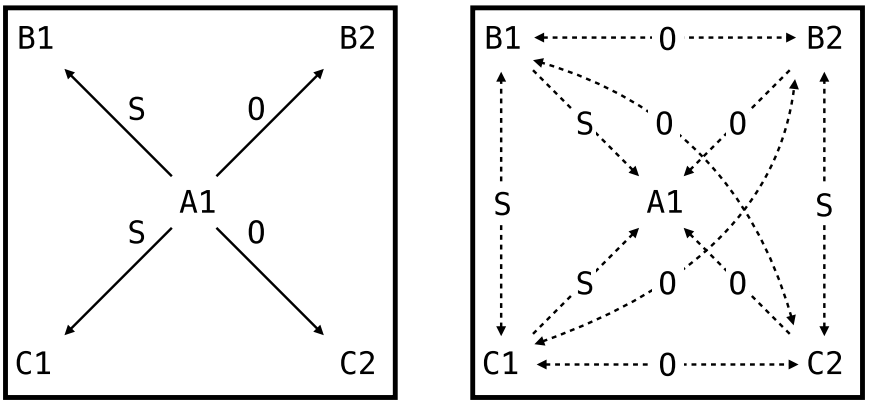


FIGURE 5
The network trained as part of the second experiment of this paper. S and O indicate SAME and OPPOSITE, respectively. Left panel shows relations that are explicitly trained. Right panel shows derived relations.

structured experiences and subsequently perform flexible relational inference provides a clear departure from contemporary AI models that primarily rely on large-scale statistical training. Instead, our approach emphasizes "small data" and logical consistency, aligning closely with the RFT premise that very few exemplars, combined with appropriate contextual cues, can generate powerful relational generalizations.

7.2 Implications for artificial general intelligence

Our theoretical demonstration of AARR within NARS offers significant implications for AGI research. First, it illustrates that sophisticated relational reasoning is achievable through adaptive symbolic systems without relying on extensive datasets, reinforcing structured symbolic learning as a viable path toward AGI. Second, our approach establishes learning psychology principles—particularly those articulated by RFT—as functional benchmarks for evaluating AGI systems' relational generalization capabilities. Third, the flexibility of NARS in dynamically constructing relational structures under uncertainty makes it suitable for adaptive, real-world contexts. Lastly, integrating adaptive logic with relational reasoning supports broad applications, including robotics and human-AI interaction, where context-sensitive symbolic manipulation is essential for achieving human-like understanding.

7.3 Limitations and future research directions

This theoretical study presents a conceptual framework and logical derivations rather than empirical validation. As such, the proposed extensions to NARS have not yet been practically implemented or empirically tested within an actual NARS-based AI system. Quantitative evaluations, such as measuring accuracy, precision, recall, or F1-score of learned relational structures, are therefore not presented in this study. Empirical validation—including quantitative performance assessments and

comparative baseline evaluations with established methods such as Inductive Logic Programming (ILP), Statistical Relational Learning (SRL), and Neural Logic Machines—remains essential future work.

Furthermore, our theoretical demonstrations employed binary (two-choice) comparisons rather than multi-choice comparison tasks typically found in human MTS studies, thereby simplifying the generalization and discrimination demands. Future empirical validations should implement multi-choice comparison setups to systematically assess the scalability and generalization of relational responding within the NARS framework.

Several other avenues remain open for further exploration. One immediate direction involves expanding the relational frames modeled in NARS beyond equivalence and opposition, including comparative, hierarchical, and deictic relations, to comprehensively evaluate the system's generalization capabilities. Another promising direction involves scaling relational networks by increasing stimulus complexity, testing NARS's resource management and inference flexibility. Additionally, integrating perceptual inputs with symbolic reasoning represents a crucial step toward practical, embodied applications, enabling NARS to generate and reason about relations directly from sensory data in dynamic environments. Lastly, further refining and automating the relational learning mechanisms within NARS, alongside comparisons of NARS-derived relational learning curves with empirical human data, could guide targeted enhancements and deepen our understanding of relational cognition in both artificial and biological systems.

8 Conclusion

We presented a theoretical framework demonstrating that NARS, enhanced by relational learning principles derived from Relational Frame Theory, can successfully model Arbitrarily Applicable Relational Responding—a foundational component of human cognition. This provides a concrete method for developing symbolic AI systems capable of dynamic, context-sensitive relational reasoning similar to that observed in humans. These findings represent a meaningful step toward bridging cognitive science and artificial intelligence, emphasizing that principles identified through human learning research can inform AI systems that "think" more like humans—not necessarily in brain-like structures but in the dynamic and contextually controlled use of symbolic knowledge. Continued interdisciplinary research in this direction holds considerable promise for developing flexible, adaptive, and ultimately more human-like artificial intelligence.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

RJ: Conceptualization, Formal Analysis, Funding acquisition, Methodology, Visualization, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was in part financially supported by Digital Futures through grant agreement KTH-RPROJ-0146472.

Acknowledgments

The author would like to thank Patrick Hammer and Tony Lofthouse for many valuable discussions regarding the work presented in this paper.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that Generative AI was used in the creation of this manuscript. The author verifies and takes full responsibility for the use of generative AI in the preparation of this manuscript. Generative AI was used to assist with editing for language clarity, improving readability, suggesting minor structural refinements, and formatting LaTeX elements. All conceptual content, theoretical ideas, analysis, and conclusions were independently developed by the author.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2025.1586033/full#supplementary-material

References

Barnes, D., and Hampson, P. J. (1993). Stimulus equivalence and connectionism: implications for behavior analysis and cognitive science. *Psychol. Rec.* 43, 617–638. doi:10.1007/bf03395903

Cassidy, S., Roche, B., Colbert, D., Stewart, I., and Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learn. Individ. Differ.* 47, 222–235. doi:10.1016/j.lindif.2016.03.001

Cropper, A., and Dumančić, S. (2022). Inductive logic programming at 30: a new introduction. *J. Artif. Intell. Res.* 74, 765–850. doi:10.1613/jair.1.13507

Cullinan, V. A., Barnes, D., Hampson, P. J., and Lyddy, F. (1994). A transfer of explicitly and nonexplicitly trained sequence responses through equivalence relations: an experimental demonstration and connectionist model. *Psychol. Rec.* 44, 559–585. doi:10.1007/bf03395144

De Houwer, J., and Hughes, S. (2020). *The psychology of learning: an introduction from a functional-cognitive perspective*. Cambridge, Massachusetts: MIT Press.

Devany, J. M., Hayes, S. C., and Nelson, R. O. (1986). Equivalence class formation in language-able and language-disabled children. *J. Exp. analysis Behav.* 46, 243–257. doi:10.1901/jeab.1986.46-243

Dixon, M. R., Peach, J., Daar, J. H., and Penrod, C. (2017). Teaching complex verbal operants to children with autism and establishing generalization using the peak curriculum. *J. Appl. Behav. analysis* 50, 317–331. doi:10.1002/jaba.373

Dymond, S., and Rehfeldt, R. A. (2000). Understanding complex behavior: the transformation of stimulus functions. *Behav. Analyst* 23, 239–254. doi:10.1007/bf03392013

Edwards, D. J. (2024). A functional contextual, observer-centric, quantum mechanical, and neuro-symbolic approach to solving the alignment problem of artificial general intelligence: safe ai through intersecting computational psychological neuroscience and llm architecture for emergent theory of mind. *Front. Comput. Neurosci.* 18, 1395901. doi:10.3389/fncom.2024.1395901

Edwards, D. J., McEnteggart, C., and Barnes-Holmes, Y. (2022). A functional contextual account of background knowledge in categorization: implications for artificial general intelligence and cognitive accounts of general knowledge. *Front. Psychol.* 13, 745306. doi:10.3389/fpsyg.2022.745306

Hammer, P. (2022). Reasoning-learning systems based on non-axiomatic reasoning system theory. In: *International workshop on self-supervised learning*. Westminster: PMLR, 89–107.

Hammer, P., and Lofthouse, T. (2020). 'opennars for applications': architecture and control. In: *International conference on artificial general intelligence*. Cham: Springer. p. 193–204.

Hayes, S. C., Devany, J. M., Kohlenberg, B. S., Brownstein, A. J., and Shelby, J. (1987). Stimulus equivalence and the symbolic control of behavior. *Rev. Mex. Anal. Conducta* 13, 361–374.

Hayes, S. C., Barnes-Holmes, D., and Roche, B. (2001). *Relational frame theory: a post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.

Hayes, S. C., Law, S., Assemi, K., Falletta-Cowden, N., Shamblin, M., Burleigh, K., et al. (2021). Relating is an operant: a fly over of 35 years of rft research. *Perspect. em Análise do Comportamento* 12, 005–032. doi:10.18761/PAC.2021.v12.RFT.02

Hughes, S., De Houwer, J., and Perugini, M. (2016). Expanding the boundaries of evaluative learning research: how intersecting regularities shape our likes and dislikes. *J. Exp. Psychol. General* 145, 731–754. doi:10.1037/xge0000100

Johansson, R. (2024a). *Empirical studies in machine psychology*. Sweden: Linköping University.

Johansson, R. (2024b). Machine psychology: integrating operant conditioning with the non-axiomatic reasoning system for advancing artificial general intelligence research. *Front. Robotics AI* 11, 1440631. doi:10.3389/frobt.2024.1440631

Johansson, R., Lofthouse, T., and Hammer, P. (2023). Generalized identity matching in nars. In: Artificial general intelligence: 15th international conference, AGI 2022; 2022 August 19–22; Seattle, WA, USA. Cham: Springer. p. 243–249.

Johansson, R., Hammer, P., and Lofthouse, T. (2024). Functional equivalence with nars. *arXiv preprint*. arXiv:2405.03340

Lenat, D. B. (1995). Cyc: a large-scale investment in knowledge infrastructure. *Commun. ACM* 38, 33–38. doi:10.1145/219717.219745

Luciano, C., Becerra, I. G., and Valverde, M. R. (2007). The role of multiple-exemplar training and naming in establishing derived equivalence in an infant. *J. Exp. Analysis Behav.* 87, 349–365. doi:10.1901/jeab.2007.08-06

Marra, G., Dumančić, S., Manhaeve, R., and De Raedt, L. (2024). From statistical relational to neurosymbolic artificial intelligence: a survey. *Artif. Intell.* 328, 104062. doi:10.1016/j.artint.2023.104062

Meli, D., and Fiorini, P. (2025). Inductive learning of robot task knowledge from raw data and online expert feedback. *Mach. Learn.* 114, 91. doi:10.1007/s10994-024-06636-6

Nitti, D., De Laet, T., and De Raedt, L. (2016). Probabilistic logic programming for hybrid relational domains. *Mach. Learn.* 103, 407–449. doi:10.1007/s10994-016-5558-8

Roche, B., Barnes-Holmes, D., Barnes-Holmes, Y., Smeets, P. M., and McGeady, S. (2000). Contextual control over the derived transformation of discriminative and sexual arousal functions. *Psychol. Rec.* 50, 267–291. doi:10.1007/bf03395356

Rosenbloom, P. S., Demski, A., and Ustun, V. (2016). The sigma cognitive architecture and system: towards functionally elegant grand unification. *J. Artif. General Intell.* 7 (1), 1–103. doi:10.1515/jagi-2016-0001

Santoro, A., Raposo, D., Barrett, D. G., Malinowski, M., Pascanu, R., Battaglia, P., et al. (2017). A simple neural network module for relational reasoning. *Adv. neural Inf. Process. Syst.* 30. doi:10.48550/arXiv.1706.01427

Schusterman, R. J., and Kastak, D. (1993). A California sea lion (zalophus californianus) is capable of forming equivalence relations. *Psychol. Rec.* 43, 823–839. doi:10.1007/bf03395915

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., and Goodman, N. D. (2011). How to grow a mind: statistics, structure, and abstraction. *science* 331, 1279–1285. doi:10.1126/science.1192788

Tovar, A. E., Torres-Chávez, Á., Mofrad, A. A., and Arntzen, E. (2023). Computational models of stimulus equivalence: an intersection for the study of symbolic behavior. *J. Exp. Analysis Behav.* 119, 407–425. doi:10.1002/jeab.829

Wang, P. (2013). *Non-axiomatic logic: a model of intelligent reasoning*. River Edge, NJ: World Scientific.

Wang, P. (2019). On defining artificial intelligence. *J. Artif. General Intell.* 10, 1–37. doi:10.2478/jagi-2019-0002

Wang, P. (2022). Intelligence: from definition to design. In: *International workshop on self-supervised learning*. Westminster: PMLR. p. 35–47.

Wang, P., Hahm, C., and Hammer, P. (2022). A model of unified perception and cognition. *Front. Artif. Intell.* 5, 806403. doi:10.3389/frai.2022.806403

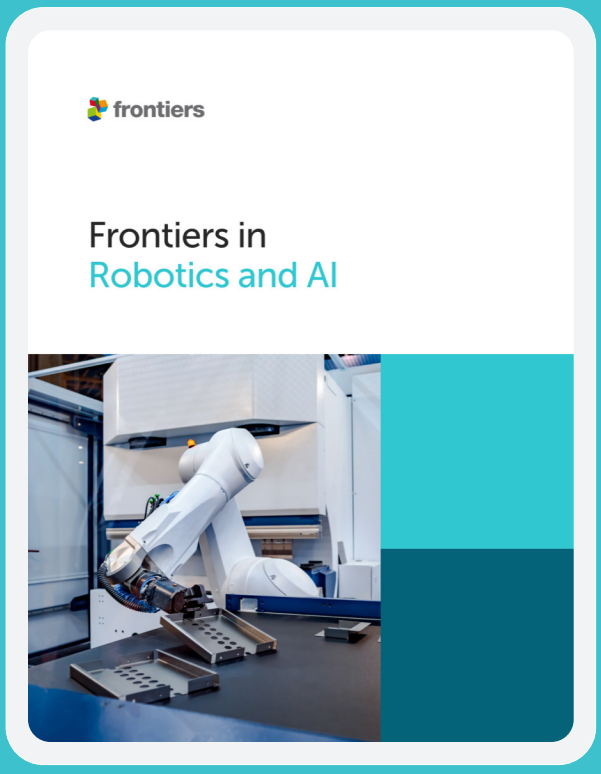
Frontiers in
Robotics and AI

Explores the applications of robotics technology for modern society

A multidisciplinary journal focusing on the theory of robotics, technology, and artificial intelligence, and their applications - from biomedical to space robotics.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact