COMMENTARY

Are Transformers Truly Foundational for Robotics?

James A. R. Marshall[1], Andrew B. Barron[2]

1 Centre for Machine Intelligence and Department of Computer Science, University of Sheffield, United Kingdom
2 School of Natural Sciences, Macquarie University, Australia

Abstract

Generative Pre-Trained Transformers (GPTs) are hyped to revolutionize robotics. Here we question their utility. GPTs for autonomous robotics demand enormous and costly compute, excessive training times and (often) offboard wireless control. We contrast GPT state of the art with how tiny insect brains have achieved robust autonomy with none of these constraints. We highlight lessons that can be learned from biology to enhance the utility of GPTs in robotics.

Corresponding author: james.marshall@sheffield.ac.uk

Introduction

Recent years have seen major advances in generative Artificial Intelligence due to the development and deployment of a new architecture; the Generative Pre-Trained Transformer (GPT), or transformer for short[1]. Through adding an attentional mechanism to deep neural networks and deploying on internet-scale training sets, transformers have led to rapid advancements in Large Language Models (LLMs) for natural language processing and generation. Following these early applications, transformers have, alongside other architectures such as diffusion models[2] been applied to the development of Visual Language Models for text to image and video (e.g.[3]) as well as other multimodal applications (e.g.[4]). These successes have inspired the investigation of transformer architectures for the robotics domain. The challenges of unstructured multimodal inputs sensed in complicated environments, coupled with high degrees of freedom in robot control, have constrained the development of robots that are simultaneously generally capable, and robust, in their behaviour. The promise of transformers for robotics appears to be that large-scale training can, through specialisation on further smaller-scale training sets, provide general and adaptable solutions to a wide variety of robotics tasks[5]. Because they can be applied across so many application domains transformer-based approaches have been labelled Foundation Models[5] indicating their supposed fundamental status but also their incomplete nature. Applications of foundation models to robotics have recently taken off in the minds of developers and researchers.

Transformers have their genesis in large language modelling (LLM). LLMs have also proved to be generalizable and transformative to many applications, but they are not without limitations. As we review below, there are increasingly

recognised issues with LLMs in the areas of training dataset size, compute resources for training, the financial and ecological costs of both, as well as robustness of behavioural output. In this article we question whether transformer architectures are likely to be truly foundational for robotics. We ask whether transformers provide the only or best route towards Artificial General Autonomy, proposing that, unlike 'intelligence',[6] the level of autonomy of a robotics system is well-defined, measurable, and economically meaningful.

Drawing on earlier critiques of GPTs and related approaches, we argue that transformers provide a facsimile of autonomy rather than true autonomy. We then review alternative approaches that have been proposed. The contrast between GPT solutions to autonomous robotics and biological solutions to autonomous behavioural control achieved by animal brains is stark. We explore this contrast to propose what is missing from current GPT approaches, and what could be added in to enhance robust and scalable robot autonomy.

Progress in Applying Transformer Architectures to Autonomy

Transformers have seen rapid application to robot autonomy. As well as high profile commercial announcements and demonstrations, end-to-end solutions to robot autonomy have been developed in the peer-reviewed literature by both academic and industrial groups, to tasks particularly focussing on robot navigation and dexterity (for a review, see[7]).

While the early promise of transformers for robot autonomy seems to be being realised, for a general and scalable solution it is essential to recognise that this technology still comes with significant limitations that will constrain future performance and adoption. While some of these may become less acute as the traditional efficiencies associated with the development and deployment of a novel technology are realised, we argue that there are fundamental structural issues with current transformer architectures, and that these should motivate a longer term search for alternative and complementary approaches, which we review later in this article.

Training Data Size and Cost Requirements are Likely to Grow

At the heart of the transformer approach to any problem is a scaling requirement. Given the lack of inductive biases these learning systems are highly flexible, however the corollary of this is that their training data requirements are vast. The usual model for deployment of a transformer-based foundation model is to train on an internet-scale corpus so that the model acquires multi-modal correspondences and domain knowledge, then further specialise on a smaller training data set for a specific set of tasks. The costs of this are very substantial. Even excluding environmental impacts, state-of-the-art LLMs cost on the order of $10s to $100s of millions per training episode[8]. For robotics applications, further training for particular tasks such as navigation and manipulation is usually required. The availability and cost of acquiring good training datasets is recognised as a major problem. Proposed solutions include the curation of open datasets covering multiple tasks and robot types[9], although currently these can be biased to a relatively small number of tasks.

There is also the extensive use of physics-based simulators to generate training data (e.g.[10]). We argue that, similarly to LLMs, exponentially increasing quantities of data are likely to be required to sustain advances in performance[11]. Even for

text and multimodal datasets where the internet provides a very large corpus of training data' for free', the availability of training data risks becoming a limiting factor[12]. For robotics datasets the costs of collecting useful training data, either physically or through simulation, will be much more acute. Furthermore, since improvement in transformers' performance is predicated on increases in scale of training data and weights this problem will only get worse.

Compute and Infrastructure Costs and Requirements will Persist

Once the costs of training a transformer-based architecture are paid, the inference costs at deployment can still be substantial. For example, Meta's Llama 3.1 has cloud-scale deployments[13] (405bn double-precision parameters). There are also reduced size and precision versions suitable for deployment on local GPUs (e.g. 8bn half-precision integer parameters), which can take ~20- 100GB of memory for inference. This demands a substantial GPU for even the simplest models being run on a robot[13]. While binarisation, quantisation, and other approaches have been used to help design edge AI accelerators for deep and convolutional neural networks[14], the scale of the problem for transformers is many orders of magnitude larger. For example, for one of the longest researched applications of deep nets, object detection, one state of the art algorithm has on the order of 10m-80m network weights[13], compared to the 8bn-405bn weights mentioned above for a state-of-the-art LLM. This represents a four orders of magnitude difference in scale, even before the additional requirements of training a transformer for robotics tasks are taken into account. Hence there is very active research into methods to avoid the cloud compute bottleneck, including utilisation of novel technologies such as 6G[15]. Moore's law and the advent of novel parallel compute architectures has traditionally saved AI, and computer software more generally. For foundation models, however, we argue that although available compute can be scaled exponentially, the exponential requirements for model size and throughput will be in opposition. A real-terms reduction in requirements for compute as performance improvements are sought will only occur when the exponent for the former is greater than the exponent for the latter. However increasingly we are considered to be in a post-Moore's Law world where further innovation in materials is required to make progress [e.g.[16]].

Hallucinations for Transformers in Robotics May Become Acute

As a consequence of their statistical training and inference, LLMs are prone to confabulation and hallucination[17][18]. While such outputs can still be damaging even for a disembodied AI, when a transformer architecture is embodied the risks are magnified. As with humans, hallucinations may manifest in ways likely to cause harm to the robot or to others, and adversarial attacks on guardrails for transformers in robotics have already been demonstrated[19]. This is likely to require that humans remain in the control loop as teleoperators to ensure robots are remotely supervised, or that robots are isolated from humans, or both. Any of these outcomes will of course limit the promised benefits of robotics. As other researchers have argued, these structural issues with statistical approaches to AI are unlikely to find remedy without significant architectural change[20].

Transformers Give a Facsimile of Intelligent Autonomy

Given the above concerns, why are transformers seeing increasing adoption for robotics? We attribute this to two factors: first, as with LLMs and VLMs, striking early advances have been made in traditionally very difficult areas, such as humanoid control, manipulation, and, of course, natural language interfaces. Second, however, we believe a tendency of human observers to anthropomorphise often leads some of them to ascribe abilities, and the potential for understanding, that the architecture does not, and cannot, technically support.

While there are many types of transformer the central motif is a repeating unit composed of a self- attention block followed by a multilayer perceptron block[21] (Figure 1, right). The control flow is feedforward, while the attention mechanism learns which earlier elements of the input to attend to in predicting the next appropriate action. As with LLMs, both the power and generalizability of transformers for robotics comes from their extensive training so that, once trained, they can perform the operation of matching an input to a predicted output. In robotics transformers succeed in resolving and executing an action from an input, but this is achieved by interpolation and extrapolation of the training set, with unreliable off-training-set performance[22]. There is no reasoning and no reason why a transformer selects one response over another, other than the selected option carrying the highest predictive weight following training[23]. The same can be said of the language abilities of LLMs, which have been described as stochastic parrots[24].

Training and reference to learned experience is an important part of biological autonomous decision making too, but for humans and other animals decision making is also supported by reasoning from models of how the world works, how other involved agents should operate, and why the selected action is situation appropriate[25]. Transformers lack these models[26][27]. An autonomous robot's capacity will be limited by the scope of the training dataset. Since transformers responses are unreasoned products of the training data, any transformer-based application cannot justify a decision other than by statistical association. This poses serious challenges for any form of human / robot interaction. If we were to ask a well-intentioned human coworker why they made an error they would do their best to explain the reasoning behind their actions. If we ask a transformer based robot why it made an error there would be no reasoned answer *per se*; the answer to the query will have at best a correlation but no causal relationship to the error made.

Alternatives and Complements to Transformers for Autonomy

If transformers are not the full answer, what is? Here we review the main alternative proposals, with an emphasis on our preferred approach, drawing deep inspiration from how the biological brain solves the autonomy problem.

Natural Intelligence

The gulf between transformer approaches to robotics and how biological brains produce autonomous behaviour is stark (Figure 1). Most often comparisons are drawn between LLMs, GPTs and human reasoning[23][25][28], but the comparison with animal brains and animal reasoning is even more pronounced. For example, the honey bee brain is tiny (just over one cubic millimeter) and contains fewer than one million neurons[29]. The number of synapses in the bee brain is not known, but if we can infer from the *Drosophila* connectome[30] there will be fewer than half a billion synapses in the bee

brain. (Figure 1, left). Demonstrably, this is all a bee needs to reliably navigate over long (several kilometre) distances, autonomously harvest pollen and nectar from the environment, communicate and coordinate their efforts with their hive mates, and perform all the many jobs needed to build and maintain their colony, including raising the next generation. They can solve complex foraging economics problems, majoring on the resources their colony needs and harvesting them from cryptic and ephemeral flowers patchily distributed in the environment[31]. Bees are able to fly with no practice, and just twenty minutes of structured flight time around the hive is enough for them to be able to navigate proficiently in their environment[32]. The contrast with the prolonged training needed by transformers could not be greater. The power consumption of a bee brain as it performs entirely on-board autonomous decision making is infinitesimal compared to any GPT. In contrast to transformers, animal brains have been massively 'pre- trained' on a planetary scale, to use minimal information and generate a very wide variety of behaviours (Figure 2).
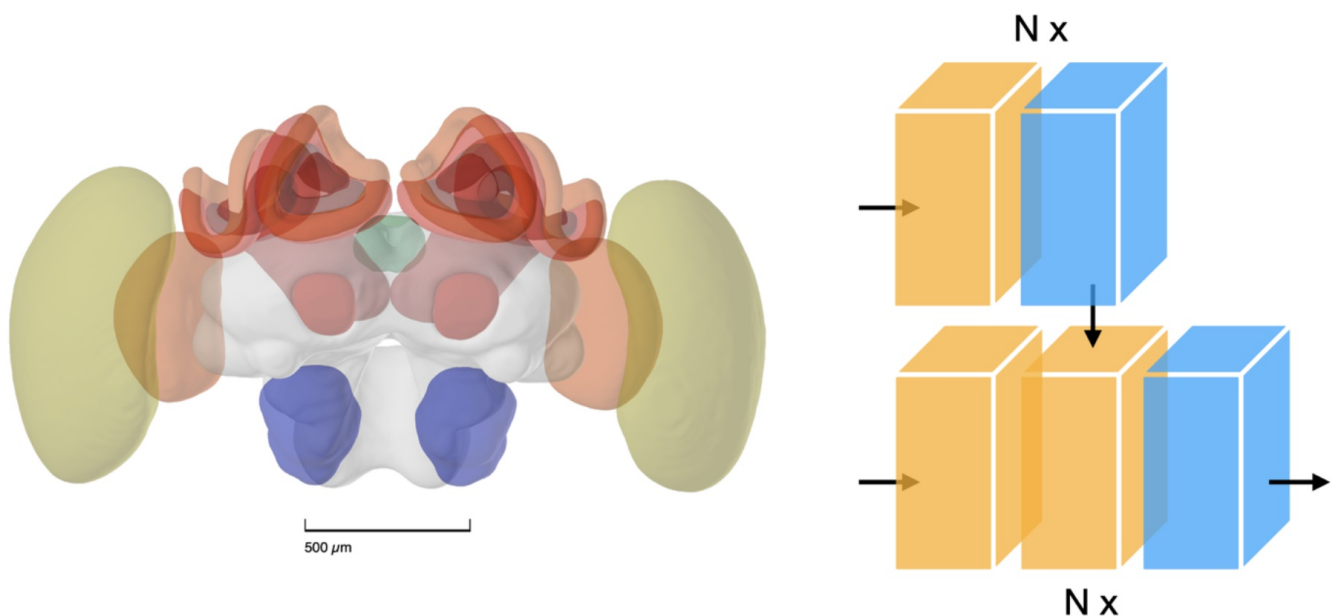

Figure 1. *Left: The brain of a honey bee forager Apis mellifera provides high levels of autonomy integrating multi-modal sensory data to navigate, communicate locations in space, learn associations between stimuli and rewards, using fewer than 960,000 neurons. Distinct brain regions specialise in perception including vision (optic lopes, yellow and orange), olfaction (antennal lobes, blue), and feed into memory centres (mushroom bodies, red). Sensory and memory pathways converge in the central complex which integrates sensing and learned associations in a single representation of the bee situated relative to percepts weighted by the bees internal state. This is sufficient to resolve competing goals, which drives behaviour directly by interfacing with premotor neurons (not shown). Brain regions are highly differentiated in structure and function according to task demands, and come together in a modular architecture with high degrees of intra-module connectivity but limited and well defined inter-module connections. Image source: insectbraindb.org Right: the generative pre-trained transformer (GPT) architecture.*
Multimodal sensory inputs are embedded in high dimensional space (not shown) then feed into repeated blocks of attentional mechanisms (yellow) followed by feed forward deep networks (blue), with intermediate normalisation and selection layers (not shown). Each block is hence a very large and non-sparse matrix, with matrix multiplications propagating through the GPT to produce the next output in sequence. Knowledge of the task is encoded in the learned values within the matrices, whose total entry counts typically range in the billions to trillions. Thus although total GPT parameters vastly exceed the number of synaptic connections in a simple brain, they are far less robust in behavioural output. Figure adapted from[1].

How does the humble bee outperform transformers in compute, energetic cost, and training time? In a word – structure. The generalisability of transformers, and arguably their elegance, is because before pre-training they are not structurally differentiated according to function. The insect brain, by contrast is a case study in structure-function specialization. The insect brain is subdivided into modules (Figure 1, left). Each module is specialised for processing different domains of the autonomous decision-making challenge. Each specialization in each module exploits the regularities and properties of the information it is processing to reduce compute and increase overall system efficiency. For example, specialized modules in the bee, ant and fly brain process the pattern of polarized light in the sky generated around the sun[33][34]. This is a valuable and robust navigational cue. Its structure is preserved by a topographic processor – the protocerebral bridge in the central complex – which outputs to a region that operates as a ring attractor to establish orientation of the animal relative to external cues[33][35][36]. This connects to yet another module which is topographically structured as the azimuth, and can support the relative localization of the insect to external objects[33][34]. The regularities of the external world are reflected in how they are represented in the insect brain, which conveys a form of intuitive physics (albeit very different from the type of physics engines used in AI). Olfactory and visual sensory lobes are each specialised to the input properties of their sensory domain. The sensory lobes sharpen, enhance and ultimately compress sensory signals for projection to multimodal sensory integration regions[37]. The largest of these, the mushroom body, has a structure similar to a three-layer neural network with an expanded middle layer[38][39][40]. This seems especially adept at multimodal classification.
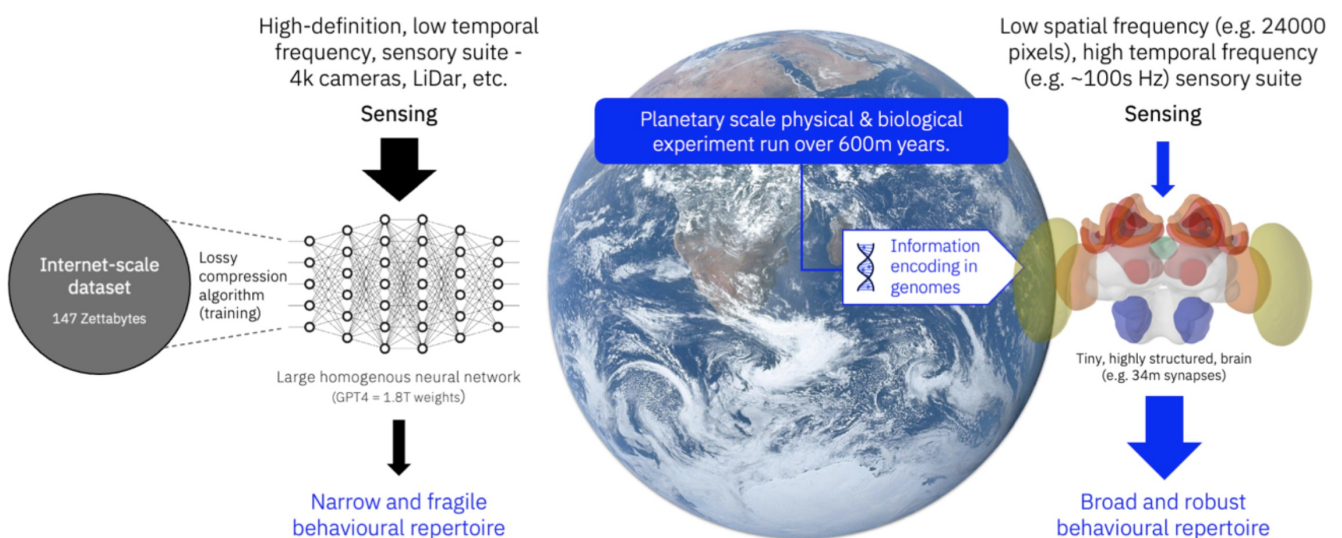

Figure 2. *Transformer-based approaches to autonomy rely on internet-scale datasets as input, which trains to process input from a suite of high resolution sensors, such as 4k cameras and LiDar, in order to provide a limited behavioural repertoire, in comparison to biological autonomous agents. In contrast, 600m years of evolution on a planetary scale, with complex physics, has encoded blueprints to build autonomous brains into the genome of a massive variety of animal species. These brains process much sparser input from specialized sensor suites, to generate a hugely rich variety of adaptive behaviours. Image sources: NASA (Earth), insectbraindb.org (honeybee brain).*

Insects lack the declarative reasoning of humans, but their reasoning is built around a form of elementary world model. Insects possess a unitary and coherent representation of external space within which they have a first-person perspective

on objects around them[41]. The valence of objects is influenced by the insect's learned experience with them, as well as innate valence and subjective physiological state[37]. Differences in valence and location of objects arbitrate the insect's selection[42][43][44]. This form of reasoning might be elementary, but it is still more comprehensible and explicit than the reasonless transformers. It is increasingly recognized that AI stands to benefit tremendously from importing concepts and algorithms from insect neuroscience[45][46].

Objective AI and World Models

Other researchers have proposed that indeed the autonomy abilities of animals (including those 'simpler than humans) should provide inspiration for AI researchers[47]. However, this inspiration is much looser than the Natural Intelligence approach above. While the 'objective AI approach does indeed propose modular AI architectures that correspond with an understanding of the human brain developed in neuroscience, cognitive science, and psychology, the proposal is actually quite different; rather than directly seek to reverse-engineer neural circuits in specialist brain modules, instead the idea is to design trainable modules that interface with each other in order to generate more adaptive behaviour than a largely undifferentiated large neural net could be expected to.

Thus, for example, rather than directly seek to understand how feature detectors in the early primate visual system function, a feature detector module would be trained. A key part of the proposal is the reintroduction of explicit and configurable worlds models, drawing inspiration from cognitive science; however these also remain trained from data[48].

Hybrid Approaches

Still other researchers, drawing on a long running proposal but also gaining renewed motivation from contemporary developments in AI, have proposed the 'neurosymbolic approach'[20]. This approach argues that, while deep nets are very suitable for perceptual tasks such as object detection, they are fundamentally unsuited to the symbolic manipulation that is part of reasoning, planning, and decision making. In the context of transformers, this has recently been vindicated by observations that LLMs fail to robustly deal with and manipulate symbolic knowledge[26][27]. Thus the proposal is to combine the perceptual strengths of statistical AI with the causal strengths of the older, symbolic, approach to AI. Given the neural bases of symbolic reasoning in the brain are poorly understood, this is a particularly pragmatic approach. In doing so it is hoped that the limitations of the first, symbolic, wave of AI, will be ameliorated by working around the problems they suffered in having sole responsibility for dealing with the perceptual complexity of the real world[49]. We suggest that an even more powerful combination could include the use of Natural Intelligence approaches to perception and modelling of space and decision option sets within it.

Conclusion

Transformer architectures have brought to robotics the rapid progress that they had already brought to natural language and muli-modal AI. However, there are reasons to continue the search for solutions to the robotics autonomy problem.

Transformer architectures treat the world in purely statistical terms, albeit grounded in perceptual inputs. This was arguably a deliberate choice in response to the' bitter lesson'[50], that inductive biases in AI have historically failed[49]. However, this results in an autonomy solution very different to the way the only truly autonomous artefact known to humanity, the biological brain, functions. Here we have highlighted this, and conclude by arguing that the tremendous recent advances in data on, and understanding of, a variety of brains, means the time is ripe to revisit the 'bitter lesson', and see what new lessons for AI can be learned from their study.

Statements and Declarations

Acknowledgements

Author Contributions

JARM and ABB conceived of, wrote, and edited the paper.

Conflicts of Interest

JARM is Founder Science Officer at and shareholder in Opteran Technologies Ltd. ABB is an unremunerated Academic Advisory Board member at Opteran Technologies Ltd.

References

1. [a, b]*Vaswani A, et al. in Proceedings of the 31st International Conference on Neural Information Processing Systems 6000–6010 (Curran Associates Inc., Long Beach, California, USA, 2017).*

2. [^]*Croitoru FA, Hondru V, Ionescu RT, Shah M. "Diffusion Models in Vision: A Survey". IEEE Transactions on Pattern Analysis and Machine Intelligence. 45: 10850-10869. doi:10.1109/TPAMI.2023.3261988.*

3. [^]*Wang Z, et al. SimVLM: Simple Visual Language Model Pretraining with Weak Supervision. arXiv:2108.10904v3, doi: 10.48550/arXiv.2108.10904 (2022).*

4. [^]*Tsai YH, et al. "Multimodal Transformer for Unaligned Multimodal Language Sequences". Proc Conf Assoc Comput Linguist Meet 2019, 6558-6569. doi:10.18653/v1/p19-1656.*

5. [a, b]Bommasani R, et al. On the Opportunities and Risks of Foundation Models. ArXiv abs/2108.07258 (2021).

6. [^]Mitchell M. "Debates on the nature of artificial general intelligence". Science. 383: eado7069. doi:10.1126/science.ado7069.

7. [^]Hu Y, et al. Toward General-Purpose Robots via Foundation Models: A Survey and Meta- Analysis. arXiv:2312.08782 (2023).

8. [^]Buchholz K. The Extreme Cost Of Training AI Models. Forbes (2024).

9. [^]Khazatsky A, et al. DROID: A Large-Scale In-The-Wild Robot Manipulation Dataset. arXiv:2403.12945 (2024).

10. [^]Makoviychuk V, et al. Isaac Gym: High Performance GPU-Based Physics Simulation For Robot Learning. arXiv:2108.10470 (2021).

11. [^]Hoffmann J, et al. Training Compute-Optimal Large Language Models. arXiv:2203.15556 (2022).

12. [^]Villalobos P, et al. Will we run out of data? Limits of LLM scaling based on human- generated data. ArXiv, doi: https://arxiv.org/abs/2211.04325 (2024).

13. [a, b, c]Schmid P, et al. Llama 3.1 - 405B, 70B & 8B with multilinguality and long context. (2024).

14. [^]Hubara I, et al. in Advances in Neural Information Processing Systems Vol. 29 (eds D. Lee et al.) (Curran Associates, Inc., 2016).

15. [^]Qu G, et al. Mobile Edge Intelligence for Large Language Models: A Contemporary Survey. arXiv:2407.18921 (2024).

16. [^]Kim KS, et al. "The future of two-dimensional semiconductors beyond Moore's law". Nature Nanotechnology. 19: 895-906. doi:10.1038/s41565-024-01695-1.

17. [^]Li Y, et al. Evaluating Object Hallucination in Large Vision-Language Models. arXiv:2305.10355 (2023).

18. [^]Zhang Y, et al. Siren's Song in the AI Ocean: A Survey on Hallucination in Large Language Models. arXiv:2309.01219 (2023).

19. [^]Robey A, et al. Jailbreaking LLM- Controlled Robots. arXiv:2410.13691 (2024).

20. [a, b]Marcus G. The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence. arXiv:2002.06177 (2020).

21. [^]Dufter P, Schmitt M, Schütze H. "Position Information in Transformers: An Overview". Computational Linguistics. 48: 733-763. doi:10.1162/coli_a_00445.

22. [^]Bommasani R, et al. (2022).

23. [a, b]Milliere R, Buckner C. A Philosophical Introduction to Language Models - Part I: Continuity With Classic Debates. ArXiv abs/2401.03910 (2024).

24. [^]Bender EM, Gebru T, McMillan-Major A, Mitchell M. in Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency 610–623 (Association for Computing Machinery, Virtual Event, Canada, 2021).

25. [a, b]Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ. "Building machines that learn and think like people". Behav. Brain Sci. 40: e253 (2017).

26. [a, b]Momennejad I, et al. Evaluating Cognitive Maps and Planning in Large Language Models with CogEval. arXiv:2309.15129 (2023).

27. [a, b]Mirzadeh I, et al. GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language

Models. arXiv:2410.05229 (2024).

28. ^Momennejad I. "A rubric for human-like agents and NeuroAI". Phil. Trans. Roy. Soc B. 378: 20210446. doi:10.1098/rstb.2021.0446.

29. ^Perry CJ, Barron AB, Chittka L. "The frontiers of insect cognition". Current Opinion in Behavioral Sciences. 16: 111-118. doi:10.1016/j.cobeha.2017.05.011.

30. ^Dorkenwald S, et al. "Neuronal wiring diagram of an adult brain". Nature. 634: 124-138. doi:10.1038/s41586-024-07558-y.

31. ^Seeley, T. D. The wisdom of the hive: the social physiology of honey bee colonies. Harvard University Press; 1995.

32. ^Capaldi, E. A., et al. Ontogeny of orientation flight in the honeybee revealed by harmonic radar. Nature. 2000;403:537-540.

33. a, b, cPfeiffer, K., Homberg, U. Organization and functional roles of the central complex in the insect brain. Annu Rev Entomol. 2014;59:165–184.

34. a, bPlath, J. A., Barron, A. B. Current progress in understanding the functions of the insect central complex. Current Opinion in Insect Science. 2015;12:11-18.

35. ^Cope, A. J., Sabo, C., Vasilaki, E., Barron, A. B., Marshall, J. A. R. A computational model of the integration of landmarks and motion in the insect central complex. PLOS ONE. 2017;12(2):e0172325. doi:10.1371/journal.pone.0172325.

36. ^Turner-Evans, D. B., Jayaraman, V. The insect central complex. Curr. Biol. 2016;26:R445–R460.

37. a, bGalizia, C. G. Olfactory coding in the insect brain: data and conjectures. European Journal of Neuroscience. 2014;39:1784-1795.

38. ^Huerta, R., Nowotny, T., Garcia-Sanchez, M., Abarbanel, H. D. L., Rabinovich, M. I. Learning classification in the olfactory system of insects. Neural Computation. 2004;16:1601-1640.

39. ^Smith, B. H., Huerta, R., Bazhenov, M., Sinakevitch, I. in Honeybee Neurobiology and Behavior (eds Giovanni C. Galizia, Dorothea Eisenhardt, Martin Giurfa) 393-409 (Springer, 2012).

40. ^Bazhenov, M., Huerta, R., Smith, B. A computational framework for understanding decision making through integration of basic learning rules. Journal of Neuroscience. 2013;33:5686-5697.

41. ^Barron, A. B., Klein, C. What insects can tell us about the origins of consciousness. Proc. Nat. Acad. Sci. USA. 2016;113:4900-4908.

42. ^Krashes, M. J., et al. A neural circuit mechanism integrating motivational state with memory expression in Drosophila. Cell. 2009;139:416-427.

43. ^Burke, C. J., et al. Layered reward signalling through octopamine and dopamine in Drosophila. Nature. 2012;492:433-437. doi:10.1038/nature11614.

44. ^Tsao, C.-H., Chen, C.-C., Lin, C.-H., Yang, H.-Y., Lin, S. Drosophila mushroom bodies integrate hunger and satiety signals to control innate food-seeking behavior. eLife. 2018;7:e35264. doi:10.7554/eLife.35264.

45. ^Webb, B. Robots with insect brains. Science. 2020;368:244-245. doi:10.1126/science.aaz6869.

46. ^de Croon, G. C. H. E., Dupeyroux, J. J. G., Fuller, S. B., Marshall, J. A. R. Insect-inspired AI for autonomous robots. *Science Robotics. 2022;7:eabl6334. doi:10.1126/scirobotics.abl6334.*

47. ^*LeCun, Y. A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27. Open Review. 2022;1:1-62.*

48. ^*Jyothir, S. V., Jalagam, S., LeCun, Y., Sobal, V. Gradient-based Planning with World Models. arXiv:2312.17227 (2023).*

49. a, b*Summerfield, C. Natural General Intelligence How understanding the brain can help us build AI. 352 (Oxford University Press, 2023).*

50. ^*Sutton, R. S., Barto, A. G. Reinforcement learning: An introduction. (MIT press, 2018).*