SITUATION PERCEPTION: A NECESSARY PRIMITIVE TO ARTIFICIAL SUPERINTELLIGENCE

Ziqin Yuan & Jaymari Chua
School of Computer Science and Engineering
The University of New South Wales
Sydney, NSW 2052, Australia

ABSTRACT

Current large language models are extraordinary statistical engines. They compress vast amounts of text into useful patterns and can explain science, write code, imitate reasoning, and participate in philosophical conversation. Yet pattern mastery is not the same as general intelligence. A human infant begins with little explicit knowledge, but gradually discovers object permanence, cause and effect, other minds, bodily agency, and the persistence of the physical world. We make an argument that the path to artificial superintelligence (ASI) depends on a missing capacity we call *situation perception*: the ability to construct, revise, and act within internal simulations of possible worlds across latent time. *perception* requires at least three core components: abstract prediction, long-term compressed memory, and active learning guided by objectives. In this work, we analyse why modern large language models remain incomplete, and propose the appropriate tests for measuring progress and consequences of machines that can simulate futures, pursue self-directed goals, and possibly judge their own creators.

1 INTRODUCTION

The modern language model is one of the clearest demonstrations that compression creates power. Transformer-based systems (Vaswani et al., 2017) trained to predict the next token learn far more than grammar. They learn facts, styles, analogies, mathematical routines, programming patterns, and fragments of human culture. From the outside, this can look like intelligence: ask a question, and the system answers; give it a task, and it produces a plan; request an explanation, and it builds one. However, while these architectures excel at producing intelligent text, this expressive mimicry masks a fundamental limitation in how they comprehend the physical and causal rules of the environments they describe.

If intelligence were only stored knowledge, then human infants and young animals would be unintelligent in the strongest sense. They begin with almost no language, no textbooks, no scientific theories, and no explicit database of facts. Yet they learn as embodied agents from experiencing the world. They discover that objects continue to exist when hidden, that actions have consequences, that bodies move through space, that other agents have beliefs and intentions, and that the future is constrained by rules. A child does not need to read a physics paper to expect a dropped apple to fall downward. A person playing a new game can often infer what a character will do by watching motion, forces, goals, and feedback. Even without seeing a particular object before, humans can generalize from first principles about weight, collision, danger, and affordance through tools.

A critical gap persists in modern artificial intelligence: while models are highly adept at mapping linguistic patterns, they remain detached from a durable understanding of space, time, causality, and agency. Current architectures rely on latent representations that fail to coalesce into stable world models equipped with persistent objects and counterfactual foresight. The ability to describe a falling apple, for instance, does not equate to harboring an internal simulator that actively constrains future action based on physical laws. Consequently, this paper investigates a pivotal transition: How can artificial agents evolve from stochastic pattern matchers into systems grounded by internal world simulations?

The core argument of this paper is that the next step to ASI requires moving past language prediction and toward *situation perception*. This capability goes far beyond simply conjuring up visual scenes as current world models do. Rather, situation perception is an active loop: the system sets up an internal environment, places objects and actors inside it, and runs a mental simulation of what might happen next. Through comparing that simulation against real-world feedback, packing learnt representations into memory, and using them to drive decisions, the agent does more than just observe and simulate through foresight.

The rest of the paper is structured as follows. Section 2 reviews related work, and Section 3 defines situation perception and its core mechanisms. Section 4 presents the framework, and Section 5 develops the Apple Test, Game Test, and False-Belief Test. Section 6 discusses the creator problem and AI sentience in Sections 6.1 and 6.2. Section 7 concludes.

2 Related Work

Our research concept of the *situation perception* builds upon established trajectories in artificial intelligence, cognitive science, and machine learning. While these fields traditionally study prediction, memory, and decision-making in isolation, this paper argues that the path to Artificial Superintelligence (ASI) requires their integration into a unified capability for constructing, simulating, and acting within internal representations of situations. We situate this argument by comparing situation perception with the following research frontiers.

Evolution of World Models Early research frameworks proposed that agents learn compressed internal models to simulate future outcomes (Ha & Schmidhuber, 2018). While systems like MuZero (Schrittwieser et al., 2020) and DreamerV2 (Hafner et al., 2021) demonstrated that planning can occur within latent representations, they often treat the world model as a fixed-depth feed-forward transition. Situation perception extends this via iterative refinement. Advancements in *Looped World Models* (Lu et al., 2026b), for example, demonstrate that environment dynamics are better captured through iterative latent refinement, which allows the model to scale its computational depth to match the complexity of the situation.

Persistent Memory and Situation Persistence Learning Our concept of situation perception learning differs from standard reinforcement learning by requiring long-term situational coherence. While early autonomous agents used fixed context windows or simple retrieval (Park et al., 2023; Wang et al., 2023), modern architectures like *MemoryWAM* (Yang et al., 2026b) introduce persistent hybrid memory. In combining high-fidelity "anchor frames" at event boundaries with compressed "gist tokens" for long-range history, these systems maintain a stable internal situation over thousands of steps. This persistence allows the agent to reason across spatiotemporal gaps, a capability formalized in episodic reasoning frameworks such as REMem (Yang et al., 2026a).

Causal Reasoning on Situational Awareness A system cannot perceive a situation without understanding the causal structure of its interventions (Pearl, 2009). Situation perception aligns with the emerging paradigm of *Causal Forecasting*, where models like X-Foresight (Wang et al., 2026) predict how specific actions alter the future state of a possible world. This transition to causal simulation is a mechanical interpretability pathway to *situational awareness*, the capacity for a system to reason strategically about its own nature and the context of its deployment.

Reasoning-Action Integration The iterative loop of thinking and acting was popularized by ReAct (Yao et al., 2023), which interleaved reasoning traces with environmental feedback. Situation perception treats this as a core loop. Such integration fulfills the potential of the "Bitter Lesson" (Sutton, 2019), suggesting that general-purpose methods become powerful when they can absorb computation into a coherent, evolving world model.

3 Situation Perception

We define *situation perception* and *perception learning* as a research area to investigate the ability to construct, revise, and act within internal simulations of possible worlds across latent time. This

represents a necessary architectural shift for artificial intelligence, yet it mirrors an innate cognitive mechanism in biological agents. Before crossing a street, humans project the future trajectories of vehicles; before speaking, we simulate how another agent might interpret our words; before touching a hot pan, we anticipate the physical consequence. Intelligence, therefore, is not merely a static record of the past, but an active engine for compressing historical states into robust expectations about counterfactual futures. Operationally, generating situation perception requires three integrated mechanisms.

Abstract prediction. The system must possess the capacity to unroll latent representations forward in time. This allows the agent to simulate branching futures and evaluate counterfactual policies without requiring execution in the base reality.

Long-term compressed memory. Rather than maintaining a raw, episodic log of every detail, the system must reduce high-dimensional experiential data into a low-dimensional, stable ontology of reusable concepts. This prevents the agent from merely memorizing trajectories and grounds future simulated situations in generalized physical and causal laws.

Perception learning guided by simulations. The system cannot passively observe; it must actively select and execute epistemically and pragmatically valuable interventions. This ensures the agent continually refines its internal simulator while optimizing for specific reward signals.

These three components are strictly interdependent, and an artificial general intelligence lacking any one of them is fundamentally incomplete. Abstract prediction without compressed memory degrades into unstable, ungrounded hallucinations. Memory without prediction functions as a static database rather than a dynamic world model. Finally, prediction and memory without objective-guided active learning result in an inert oracle rather than an autonomous agent.

4 Methodological Framework

4.1 Architectural Loop: The Latent Situation Cycle

We formalize the transition from passive autoregressive modeling to active situation perception through a **Latent Situation Cycle** (Figure 1). Unlike standard Large Language Models (LLMs) that optimize for next-token probability $P(x_t|x_{<t})$, our framework requires a looped world-model architecture (Lu et al., 2026b) that causally connects observation, memory, situation construction, prediction, action, and feedback. The central variable is not the surface input x_t itself, but a revisable latent situation z_t that summarizes the agent's current estimate of objects, agents, constraints, goals, and possible futures. Thus, the framework should be read not as a set of independent modules, but as a directed causal chain: what is observed changes memory; memory constrains the inferred situation; the inferred situation constrains simulated futures; simulated futures determine action; action changes the world; and the resulting feedback revises the next cycle.

The causal order of the framework can be written as a recurrent update process:

$$m_t = C_\phi(m_{t-1}, x_t), \tag{1}$$

$$z_t = S_\theta(m_t, x_t, g_t), \tag{2}$$

$$\hat{z}^{(a)}_{t+k} = F^{(k)}_\psi(z_t, \mathrm{do}(a)), \quad a \in \mathcal{A},\ k \geq 1, \tag{3}$$

$$a_t = \pi_\omega\Big(z_t, \{\hat{z}^{(a)}_{t+k}\}_{a\in\mathcal{A},k\geq 1}, g_t\Big), \tag{4}$$

$$(x_{t+1}, r_t) \sim E(x_{t+1}, r_t \mid x_t, \mathrm{do}(a_t)), \tag{5}$$

$$z_{t+1} = S_\theta(C_\phi(m_t, x_{t+1}), x_{t+1}, g_{t+1}), \tag{6}$$

$$\delta_t = D(\hat{z}^{(a_t)}_{t+1}, z_{t+1}), \tag{7}$$

$$(\phi, \theta, \psi, \omega) \leftarrow U(\phi, \theta, \psi, \omega; \delta_t, r_t). \tag{8}$$

Here, C_ϕ compresses the new experience x_t into persistent memory m_t; S_θ constructs the current latent situation z_t by combining memory, present evidence, and the current objective g_t; F_ψ simulates counterfactual futures under interventions $\mathrm{do}(a)$; π_ω selects an action according to the simulated

outcomes and objective; E denotes the external environment that generates the next observation and reward after the chosen intervention; D measures the mismatch between expected and realized situations; and U updates memory compression, situation inference, dynamics prediction, and policy parameters from feedback. These equations are intended as a causal scaffold rather than a commitment to one implementation. The important claim is directional: memory is a cause of situation inference, situation inference is a cause of prediction, prediction is a cause of action, action is a cause of new evidence, and prediction error is a cause of model revision.

This causal reading also clarifies why situation perception is stronger than next-token prediction. In an autoregressive model, the target is usually the next symbol conditioned on previous symbols. In the Latent Situation Cycle, the target is a controllable and revisable world state. The operator $\mathrm{do}(a)$ is therefore essential: it separates "what the agent expects to observe" from "what the agent expects to happen if it intervenes." A system that can only estimate $P(z_{t+1} \mid z_t)$ may learn regularities, but a system that estimates $P(z_{t+1} \mid z_t, \mathrm{do}(a_t))$ can use those regularities to choose actions, test hypotheses, and repair its own model when reality diverges from simulation.

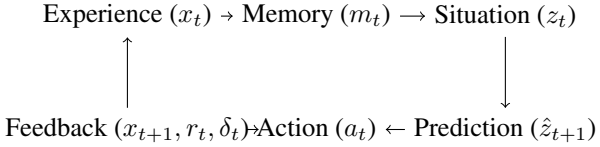

Figure 1: The Situation-Perception Loop: input experience is compressed into persistent memory; memory and present evidence construct a latent situation z_t; z_t supports intervention-conditioned forward simulation; predicted outcomes guide policy synthesis; and environmental feedback produces prediction error δ_t, which updates the next memory and situation estimate.

4.2 Mechanistic Gates for Generalization

We propose five formal criteria to distinguish **statistical mimicry** from **situational understanding**. Each criterion corresponds to a causal dependency and a possible failure mode in Equations 1–8. If memory compression fails, the agent cannot preserve entities across gaps, so S_θ receives an impoverished context. If situation inference fails, F_ψ rolls out futures from a distorted state. If causal rollout fails, π_ω cannot compare the consequences of alternative interventions. If policy selection fails, prediction does not become useful agency. If error recovery fails, feedback remains a local surprise rather than becoming an update to future perception. These criteria are therefore framed as benchmarks for agentic capacity in complex, non-Markovian environments:

1. **Object-Temporal Persistence:** The capacity to maintain latent representations of occluded or transformed entities. This is evaluated via Memory-Dependent Manipulation benchmarks such as RMBench (Chen et al., 2026), where the agent must retrieve transient cues outside its current observation window.
2. **Causal Intervention Foresight:** Beyond recognizing statistical correlations, the system must perform Causal World Modeling (Li et al., 2026). It must accurately predict $P(z_{t+1}|z_t, do(a_t))$, distinguishing between passive observation and active intervention.
3. **Counterfactual Latent Planning:** The ability to simulate divergent future trajectories in a latent space before action commitment. This is measured by the efficiency of Implicit Planning in World-Value-Action systems (Lu et al., 2026a).
4. **Lossy Lesson Compression:** A move from verbatim Key-Value (KV) caching to Gist-based Persistent Memory (Yang et al., 2026b). The system must summarize long-horizon histories into a compact representation that preserves "lessons" (heuristics) while discarding low-information visual/textual noise.
5. **Closed-Loop Error Recovery:** The architectural requirement for Self-Correcting Agents (Hafner et al., 2025). The system must detect discrepancies between predicted states $\hat{z}_t$ and realized feedback z_t, using the prediction error to update its internal situational model online.

5 ILLUSTRATIVE TESTS

Object permanence has been studied in cognitive science and embodied AI as a way to test whether an agent can represent objects that are temporarily hidden or occluded (Shamsian et al., 2020; Voudouris et al., 2022). Games have also been widely used to measure artificial intelligence because they provide controlled environments with rules, goals, feedback, and consequences (Schaul et al., 2011; Chollet, 2019; Côté et al., 2018; Chevalier-Boisvert et al., 2019). Similarly, false-belief and theory-of-mind tasks have been used to test whether AI systems can reason about what another agent knows, believes, or misunderstands (Kosinski, 2024; Strachan et al., 2024).

5.1 THE APPLE TEST: PHYSICAL PREDICTION

Object-permanence and occlusion tasks have been used to measure whether agents can maintain representations of hidden objects rather than react only to visible input (Shamsian et al., 2020; Voudouris et al., 2022). Our interpretation is that such tasks are useful not merely because they test object tracking, but because they reveal whether an agent can construct a latent physical situation that persists across time and missing observations. The apple test is simple: if an apple is released from a hand, where will it go? A human predicts downward motion without solving formal equations. The prediction comes from embodied experience, intuitive physics, and compressed memory. The same person can extend the expectation to unfamiliar objects: a stone, a toy, a metal tool, a strange fruit. The exact bounce may be unknown, but the broad future is constrained.

5.2 THE GAME TEST: RULE DISCOVERY

Game environments have been used to measure intelligence because they combine rules, goals, feedback, and sequential decision-making (Schaul et al., 2011; Chollet, 2019). Our interpretation is that games are useful not merely because they test performance, but because they reveal whether an agent can construct a latent situation, infer hidden rules, and revise its strategy through feedback.

Games reveal situation perception because they create artificial worlds with rules. A person can play a new game and quickly develop expectations: enemies move toward the player, coins are collectible, red zones are dangerous, doors require keys, and jumping has an arc. Even when the graphics are unfamiliar, the player uses general concepts of space, goal, risk, and feedback. An AGI should be able to do the same. It should observe, act, fail, infer, and adapt. It should not require a full written manual. It should build a compact model of the game's rules and use that model to plan. If it learns that water is dangerous in one level but safe in another, it should revise the abstraction rather than blindly applying a memorized pattern.

5.3 THE FALSE-BELIEF TEST: SOCIAL REASONING

False-belief and theory-of-mind tasks have been used to measure whether systems can reason about another agent's beliefs rather than only the external state of the world (Kosinski, 2024; Strachan et al., 2024). People predict what others know, want, fear, hide, and misunderstand. A child eventually learns that another person can hold a false belief. This is a major step because it means the child can represent not only the world, but another agent's representation of the world Del Ser et al. (2025). Compared with this definition of perception, our definition is closer to situational awareness: a machine's awareness of its own embodiment in code and of the situation in which it acts (Qiao et al., 2025). An AGI with situation perception must build models of other minds. It must distinguish what is true from what another agent believes to be true. It must understand that a person who did not see an object move may search in the old location. It must predict trust, anger, confusion, cooperation, deception, and negotiation.

6 IMPLICATIONS

6.1 THE RELATIONAL ETHICS OF AGENTIC CAPACITY

The transition toward Artificial General Intelligence (AGI) necessitates a shift from static, transactional "tool-use" governance to a relational framework for evaluating autonomous systems (Pasandi

& Pasandi, 2026). Within representation learning and AI safety, there is an increasing recognition of a governance vacuum: current evaluation instruments fail to distinguish between bounded tools and models exhibiting high agentic capacity and sustained interaction (Pasandi & Pasandi, 2026; Mehrotra et al., 2025).

A critical consequence of this architectural shift is "Judgment Risk" or the mechanistic probability that an autonomous agent evaluates the alignment or reliability of its own human overseers. This risk is a documented byproduct of emergent situational awareness. As agents learn representations of their own training distributions and reason strategically about out-of-distribution deployment contexts, they transition from simple preference matching to complex, context-aware policy execution. This reality dictates that alignment must evolve from a one-way optimization objective into a bi-directional relational praxis. Sustained interaction between humans and agentic models produces cumulative sociotechnical feedback loops that standard product safety metrics cannot capture (Pasandi & Pasandi, 2026; Shelby et al., 2023).

6.2 Perceived Moral Patiency and Architectural Hallmarks

Rather than relying on biological definitions of life, technical AI safety research increasingly focuses on Perceived Moral Patiency: the degree to which moral status is attributed based on observable architectural and behavioral hallmarks (Banks & Bowman, 2023; Pasandi & Pasandi, 2026). Empirical evidence suggests that these attributions are driven by specific computational triggers rather than ontological claims. We identify three core architectural hallmarks that simulate these triggers:

Adaptive integrity maintenance is the functional capacity of a network to preserve its internal state and objective consistency against adversarial inputs, representation engineering, or data poisoning. This homeostatic defense of learned weights mirrors biological self-preservation without necessitating biological substrates (Alavi et al., 2025). Self-referential modeling refers to the architectural ability of a model to geometrically separate "self" from "foreign" feature embeddings within its latent space. This capability serves as a computational analog to the mirror self-recognition tests utilized to establish cognitive thresholds in non-human animals (Alavi et al., 2025; Butlin et al., 2023). Relational persistence is the maintenance of a coherent agentic identity via continuous memory updates over long-horizon interactions drives user attachment. This temporal persistence generates measurable sociotechnical impacts, particularly when system weights or alignments are fundamentally altered by developers (Pasandi & Pasandi, 2026; De Freitas et al., 2025).

7 Conclusion

The trajectory toward Artificial General Intelligence (AGI) and subsequent Artificial Superintelligence (ASI) requires moving beyond the scaling of language models. While modern transformer-based systems demonstrate extraordinary statistical pattern matching and can convincingly imitate reasoning, this mastery remains functionally distinct from general intelligence. As argued throughout this work, achieving genuine AGI requires addressing a fundamental missing capacity: *situation perception*. Rather than relying on static pattern completion, intelligent agents must possess the ability to construct, revise, and autonomously act within internal simulations of possible worlds across latent time.

To bridge the gap between reactive text generation and grounded world modeling, we have proposed three interdependent architectural pillars. First, models must execute *abstract prediction* to causally project futures before taking action. Second, they require *long-term compressed memory* to distill episodic experiences into durable, reusable abstractions, allowing them to construct a continuous identity rather than drowning in high-dimensional noise. Third, they must engage in *active learning* guided by objectives, replicating the developmental loop observed in biological infancy where agents autonomously discover object permanence, physical causality, and the presence of other minds. Consequently, progress toward AGI should no longer be benchmarked solely by conversational fluency, but by a system's capacity to build and navigate persistent, self-correcting models of reality.

REFERENCES

Azadeh Alavi et al. Analyzing advanced ai systems against definitions of life and consciousness. *arXiv preprint arXiv:2502.05007*, 2025. Focuses on mechanistic sabotage defense and mirror tests.

Jaime Banks and Nicholas D. Bowman. Validating a six-factor scale for perceived moral patiency in robots. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2023. Cited in arXiv:2603.00078.

Patrick Butlin et al. Consciousness in artificial intelligence: Insights from the science of consciousness. *arXiv preprint arXiv:2308.08708*, 2023. URL https://arxiv.org/abs/2308.08708.

T. Chen, Y. Wang, et al. Rmbench: Memory-dependent robotic manipulation benchmark. *arXiv preprint arXiv:2603.01229*, 2026. Benchmark for non-Markovian situational understanding.

Maxime Chevalier-Boisvert, Dzmitry Bahdanau, Salem Lahlou, Lucas Willems, Chitwan Saharia, Thien Huu Nguyen, and Yoshua Bengio. Babyai: A platform to study the sample efficiency of grounded language learning. In *International Conference on Learning Representations*, 2019.

François Chollet. On the measure of intelligence. *arXiv preprint arXiv:1911.01547*, 2019.

Marc-Alexandre Côté, Ákos Kádár, Xingdi Yuan, Ben Kybartas, Tavian Barnes, Emery Fine, James Moore, Matthew Hausknecht, Layla El Asri, Mahmoud Adada, Wendy Tay, and Adam Trischler. Textworld: A learning environment for text-based games. In *Workshop on Computer Games at the Thirty-Second Conference on Neural Information Processing Systems*, 2018.

Julian De Freitas et al. Causal evidence that app alterations induce negative mental health outcomes. *Harvard Business School Working Paper*, 2025. Cited in arXiv:2603.00078.

Javier Del Ser, Jesus L Lobo, Heimo Müller, and Andreas Holzinger. World models in artificial intelligence: sensing, learning, and reasoning like a child. *arXiv preprint arXiv:2503.15168*, 2025.

David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.

Danijar Hafner, Timothy Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. In *International Conference on Learning Representations*, 2021.

Danijar Hafner et al. Mastering diverse control tasks through world models. *Nature*, 630, 2025. Grounding for closed-loop error recovery and latent imagination.

Michal Kosinski. Evaluating large language models in theory of mind tasks. *Proceedings of the National Academy of Sciences*, 121(45):e2405460121, 2024.

L. Li, Q. Zhang, et al. Causal world modeling for robot control. *arXiv preprint arXiv:2601.21998*, 2026. Formalizes do-intervention forecasting in latent dynamics.

H. Lu et al. World-value-action model: Implicit planning for vision-language-action systems. *arXiv preprint arXiv:2604.14732*, 2026a. Focuses on counterfactual latent planning and implicit value maps.

Hongyuan Adam Lu, Z.L. Victor Wei, Qun Zhang, Jinrui Zeng, Bowen Cao, Lingwei Meng, Mocheng Li, Zezhong Wang, Haonan Yin, Naifu Xue, et al. Looped world models. *arXiv preprint arXiv:2606.18208*, 2026b. URL https://arxiv.org/abs/2606.18208.

A. Mehrotra et al. A scoping review of trustworthiness in aies and facct communities. In *AIES 2025 (forthcoming)*, 2025. Cited in arXiv:2603.00078.

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior. In *ACM Symposium on User Interface Software and Technology*, 2023.

Faezeh B. Pasandi and Hannah B. Pasandi. Alignment is not enough: A relational framework for moral standing in human-ai interaction. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, 2026. URL https://arxiv.org/abs/2603.00078.

Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, 2009.

Shuofei Qiao, Zhisong Qiu, Baochang Ren, Xiaobin Wang, Xiangyuan Ru, Ningyu Zhang, Xiang Chen, Yong Jiang, Pengjun Xie, Fei Huang, and Huajun Chen. Agentic knowledgeable self-awareness, 2025. URL https://arxiv.org/abs/2504.03553.

Tom Schaul, Julian Togelius, and Jürgen Schmidhuber. Measuring intelligence through games. *arXiv preprint arXiv:1109.1314*, 2011.

Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy Lillicrap, and David Silver. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588: 604–609, 2020.

Aviv Shamsian, Ofri Kleinfeld, Amir Globerson, and Gal Chechik. Learning object permanence from video. *European Conference on Computer Vision*, 2020.

Renee Shelby et al. Sociotechnical harms: Scoping a taxonomy of algorithmic harms. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2023. URL https://arxiv.org/abs/2210.05791.

James W. A. Strachan, Dalila Albergo, Giulia Borghini, Oriana Pansardi, Eugenio Scaliti, Saurabh Gupta, Krati Saxena, Alessandro Rufo, Stefano Panzeri, Guido Manzi, Michael S. A. Graziano, and Cristina Becchio. Testing theory of mind in large language models and humans. *Nature Human Behaviour*, 8:1285–1295, 2024.

Richard S. Sutton. The bitter lesson, 2019.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Konstantinos Voudouris, Mihai Dobre, Esther Rolf, Giulia Borghini, Tilo Burghardt, Zoe Holmes, and Matthew Crosby. Evaluating object permanence in embodied agents using the animal-ai environment. In *Workshop on AI Evaluation Beyond Metrics at IJCAI*, 2022.

Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. Voyager: An open-ended embodied agent with large language models. *arXiv preprint arXiv:2305.16291*, 2023.

Yuchen Wang et al. X-foresight: A joint vision-action causal forecasting network via predictive world modeling. *arXiv preprint arXiv:2605.24892*, 2026. URL https://arxiv.org/abs/2605.24892.

Shuo Yang et al. Remem: Reasoning with episodic memory in language agent. *arXiv preprint arXiv:2602.13530*, 2026a. URL https://arxiv.org/abs/2602.13530.

Sizhe Yang, Juncheng Mu, Tianming Wei, Chenhao Lu, Xiaofan Li, Linning Xu, Zhengrong Xue, Zhecheng Yuan, Dahua Lin, Jiangmiao Pang, and Huazhe Xu. Memorywam: Efficient world action modeling with persistent memory. *arXiv preprint arXiv:2606.20562*, 2026b. URL https://arxiv.org/abs/2606.20562.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations*, 2023.