

Action-Aware World Models for Long-Horizon Planning and Decision-Making in Embodied Robotics

Aaliyah Manson

School of Electrical Engineering and Computer Science, Oregon State University, Corvallis,
OR, USA.

amanson@oregonstate.edu

Wilfrido Hernandez

Department of Computer Science, University of Houston, Houston, TX, USA.

hernandezwilfrido@uh.edu

Elodia Arrieta

Department of Computer Science, George Mason University, Fairfax, VA, USA.

arrietaelodia@gmu.edu

Abstract

Embodied robotics is moving from narrow perception-action pipelines toward systems that must reason over extended horizons under uncertainty, resource scarcity, and social accountability. World models offer a promising route by allowing agents to simulate futures, evaluate candidate actions, and update internal representations before committing physical behavior. Yet conventional world models often remain observation-centric, treating action as an external perturbation rather than as a constitutive signal that shapes memory, belief revision, and long-term control. This paper examines action-aware world models as socio-technical infrastructures for long-horizon planning and decision-making in embodied robotics. It argues that action-aware memory is not merely an architectural refinement but a systems-level requirement for maintaining coherence across exploration, interaction, adaptation, and deployment. The discussion analyzes structural trade-offs among predictive fidelity, computational cost, interpretability, latency, safety, and sustainability. It then considers governance implications, including fairness in human-robot interaction, auditability of learned policies, robustness under distribution shift, and policy frameworks for autonomous systems. Through cross-domain comparisons with reinforcement learning, probabilistic robotics, and large-scale machine learning infrastructure, the paper identifies open challenges in evaluation, memory governance, and lifecycle sustainability. The conclusion emphasizes that action-aware world models must be designed as accountable institutional systems, not only as algorithmic artifacts, if they are to support reliable long-horizon embodied decision-making.

Keywords

action-aware world models; embodied robotics; long-horizon planning; decision-making; memory architectures; robustness; governance; socio-technical systems.

1. Introduction

Embodied robotics has advanced rapidly through deep learning, reinforcement learning, and increasingly capable perception systems [1][2]. Robots can now acquire visuomotor skills, navigate unstructured environments, and manipulate objects with varying degrees of autonomy [9][10]. However, the transition from short-horizon reactive behavior to long-

horizon planning remains difficult because physical agents must maintain coherent beliefs, anticipate consequences, and adapt to changing conditions over minutes, hours, or longer. A robot that succeeds in a single task may fail when required to remember prior failures, infer hidden intentions, or coordinate with humans across multiple stages. This gap has motivated renewed interest in world models, which allow agents to learn internal representations of environment dynamics and use them for prediction, planning, and policy improvement [3][4].

Traditional world models emphasize observation prediction and latent state transitions, often treating actions as conditioning variables that influence future observations. This perspective has proven valuable in simulated control and model-based reinforcement learning, but it can underrepresent the role of action in structuring memory and belief. In embodied settings, actions are not merely outputs; they are epistemic probes that reveal hidden state, create commitments, consume resources, and alter the future space of possibilities. An action-aware world model therefore treats action as a first-class element of representation, memory, and planning. It asks not only what will happen next, but also what the agent has done, what those actions mean for future options, and how memory should be organized around interventions rather than passive observation alone.

This paper develops a system-level analysis of action-aware world models for long-horizon planning and decision-making in embodied robotics. The central claim is that action-aware memory and action-conditioned prediction are necessary but insufficient on their own; they must be embedded in an architecture that addresses latency, robustness, sustainability, fairness, and governance. The argument proceeds through conceptual foundations, architectural trade-offs, planning mechanisms, memory and state estimation, infrastructure deployment, policy implications, case comparisons, and evaluation challenges. The aim is to connect algorithmic advances with the institutional and infrastructural conditions under which embodied systems operate.

2. Conceptual Foundations of Action-Aware World Models

World models can be understood as learned or engineered structures that approximate environment dynamics for the purpose of prediction and control. In probabilistic robotics, state estimation and planning are grounded in explicit belief distributions, observation models, and transition models [20]. In reinforcement learning, model-based agents learn latent dynamics and use imagined trajectories to improve policies [2][4]. In deep learning, representational capacity enables high-dimensional perceptual inputs to be compressed into latent spaces that support prediction and generalization [1]. Action-aware world models extend these traditions by emphasizing the informational role of actions in shaping both the latent state and the memory that supports future decisions.

The distinction between observation-centric and action-aware models is consequential for long-horizon behavior. An observation-centric model may predict the next image or latent state accurately while failing to encode why a state was reached, which actions have already been attempted, and which commitments remain open. This limitation becomes acute in partially observable environments where the same observation can arise from different histories [6]. Action-aware models address this by conditioning transitions on action semantics and by maintaining memory structures that record interventions, outcomes, and unresolved hypotheses. In this sense, action awareness is closely tied to the concept of epistemic action, where movement and manipulation are used to reduce uncertainty rather than merely to achieve immediate reward.

Another foundational issue concerns the relationship between action-aware world models and representation. Classical reactive architectures argued that intelligent behavior can emerge without centralized representations [19]. Modern world models reintroduce representation but in a learned, distributed, and often opaque form. Action-aware memory offers a middle path by structuring representation around temporally extended action histories rather than static world descriptions. This structuring can improve sample efficiency, credit assignment, and transfer, but it also introduces risks of memory contamination, spurious causal associations, and overfitting to past action patterns. A robust conceptual foundation must therefore specify what counts as an action, how action histories are abstracted, and how memory is revised when the environment or task changes.

3. Architecture and System-Level Trade-offs

From an architectural perspective, action-aware world models sit at the intersection of perception, memory, prediction, planning, and control. A typical system may include perceptual encoders, latent dynamics models, action encoders, memory banks, planning modules, and safety monitors. Each component introduces trade-offs. Higher predictive fidelity may require deeper models and longer computation, which increases latency and energy consumption. Richer memory may improve long-horizon coherence but can also create retrieval bottlenecks and stale representations. Stronger action conditioning may improve controllability but can make the model sensitive to action distribution shift. These trade-offs are not merely implementation details; they shape what kinds of embodied autonomy are feasible in real-world deployments.

A central architectural tension is between reactive responsiveness and deliberative planning. Embodied robots often operate under strict timing constraints, especially when interacting with humans or moving through dynamic spaces. A world model that requires extensive rollouts before each action may be too slow, while a purely reactive policy may lack the foresight needed for long-horizon tasks. Action-aware memory can mediate this tension by caching reusable plans, action outcomes, and context-dependent affordances. A recent action-aware memory framework illustrates how interactive world models can move beyond mere exploration by organizing memory around actions and their consequences [14]. Such designs suggest that system performance depends less on raw model size than on how memory is indexed, updated, and queried under real-time constraints.

A third trade-off concerns interpretability and control. Opaque latent world models can achieve strong predictive performance but are difficult to audit, debug, or align with human expectations. Action-aware representations may improve interpretability because they connect internal states to observable action histories, yet they can also encode subtle biases in how actions are labeled or prioritized. System architects must therefore decide where to place transparency mechanisms, how to expose uncertainty to downstream planners, and how to enable human operators to override or interrogate model predictions. These decisions affect trust, liability, and operational safety, particularly in domains such as healthcare, manufacturing, transportation, and domestic assistance.

4. Long-Horizon Planning and Decision-Making

Long-horizon planning in embodied robotics requires the agent to reason across multiple timescales, from moment-to-moment control to task-level sequencing and mission-level objectives. Action-aware world models support this by enabling imagined rollouts that are conditioned on candidate action sequences. Unlike open-loop prediction, action-conditioned

imagination allows the agent to compare alternatives, estimate risk, and revise plans before acting. This capability connects to classical planning under uncertainty [6] and to modern model-based reinforcement learning [4]. However, long-horizon planning magnifies errors because small prediction biases can compound across many steps. Action-aware memory can mitigate this by grounding predictions in past action-outcome pairs, but it cannot eliminate the fundamental challenge of distribution shift.

Decision-making also involves value alignment and constraint satisfaction. A robot may need to balance task completion, safety, energy use, human comfort, and legal compliance. These objectives are often in tension and may change across contexts. Action-aware world models can represent action consequences not only in terms of reward but also in terms of resource consumption, social acceptability, and reversibility. For example, an action that opens a door may be reversible, while one that spills a liquid may not be. By encoding such properties, the model can support planning that respects constraints rather than merely maximizing a scalar objective. This perspective aligns with broader concerns in artificial intelligence safety, where specification gaming and side effects must be anticipated [13].

A further dimension is multi-agent and human-in-the-loop planning. Embodied robots rarely operate in isolation; they coordinate with humans, other robots, and infrastructure. Action-aware memory can capture interaction histories, commitments, and social norms, enabling more coherent collaboration. Yet multi-agent settings introduce nonstationarity because other agents adapt and may have hidden goals. Planning then becomes partially observable and strategically coupled [6][7]. The system must decide when to act, when to observe, when to communicate, and when to defer. These decisions are not purely algorithmic; they reflect governance choices about autonomy, responsibility, and acceptable risk.

5. Memory, State Estimation, and Robustness

Memory is the connective tissue of action-aware world models. Without memory, an agent cannot accumulate evidence, learn from past interventions, or maintain commitments over time. With memory, the agent faces problems of retrieval, compression, forgetting, and interference. Action-aware memory differs from generic sequence memory because it organizes experience around actions and their outcomes, making it possible to retrieve analogous situations and counterfactuals. This structure can improve robustness to perceptual ambiguity, since an ambiguous observation may be resolved by recalling which actions led to it and what followed. Probabilistic robotics offers mature methods for belief updating under uncertainty [20], and these can inform how action-aware memory should represent confidence and revise beliefs.

Robustness also depends on how the model handles distribution shift, sensor degradation, and adversarial conditions. A world model trained in simulation may fail in the real world because of unmodeled dynamics, lighting changes, or contact physics [11][12]. Action-aware memory can support adaptation by detecting mismatches between predicted and observed action outcomes, then updating or reweighting stored experiences. However, memory can also propagate errors if corrupted experiences are retrieved and reused. Robust systems therefore require mechanisms for uncertainty estimation, anomaly detection, memory validation, and selective forgetting. These mechanisms are not optional add-ons; they are core to safe long-horizon autonomy.

Finally, robustness has a temporal dimension. Long-horizon tasks may last long enough for components to fail, batteries to drain, or environments to change. An action-aware world

model should represent the agent's own condition, including actuator degradation, energy reserves, and maintenance needs. This self-modeling connects internal state estimation with operational planning. A robot that knows it is becoming less reliable may choose safer actions, request assistance, or defer tasks. Such self-awareness is a system-level property that emerges from the integration of memory, prediction, and governance constraints.

6. Infrastructure, Deployment, and Sustainability

Deploying action-aware world models requires infrastructure that extends far beyond the robot itself. Training may occur in cloud data centers, simulation farms, or distributed edge clusters. Inference may occur on-board, off-board, or in a hybrid configuration depending on latency, privacy, and connectivity. Data pipelines must handle multimodal sensor streams, action logs, human feedback, and environmental metadata. Model updates must be versioned, validated, and rolled back if necessary. These requirements resemble the hidden technical debt of machine learning systems, where data dependencies, configuration issues, and feedback loops create long-term maintenance burdens [17]. For embodied robotics, the stakes are higher because software errors can cause physical harm.

Sustainability is another infrastructural concern. Large world models can consume substantial energy during training and inference, and embodied robots add further costs through movement, sensing, and computation. A system that is accurate but energy-intensive may be unsuitable for mobile platforms or large fleets. Action-aware memory can improve sustainability by reducing redundant computation, enabling experience reuse, and supporting selective planning. Yet memory storage and retrieval also have costs. Designers must therefore evaluate the full lifecycle: data collection, annotation, training, deployment, maintenance, decommissioning, and recycling. Sustainability is not merely an environmental metric; it is also an economic and operational constraint that shapes which architectures can be maintained at scale.

Infrastructure also determines equity and access. If action-aware world models require expensive compute, proprietary data, and specialized engineering teams, then only well-resourced institutions can develop and deploy them. This concentration may skew research agendas, limit reproducibility, and exclude communities from shaping the technology. Open benchmarks, shared simulation environments, and modular interfaces can lower barriers, but they also raise questions about safety, misuse, and standardization. The deployment ecosystem must balance innovation with accountability, ensuring that infrastructure choices do not silently determine social outcomes.

7. Governance, Fairness, and Policy Implications

Action-aware world models raise governance questions because they inform decisions that affect people. A robot may decide how to allocate attention, when to intervene, whom to assist, and which risks to take. If the model's memory reflects historical biases, it may reproduce or amplify them. For example, action histories collected in biased environments may lead the robot to associate certain people, places, or behaviors with negative outcomes. Fairness research in machine learning has shown that bias can arise from data, objectives, and deployment contexts [16]. Governance must therefore address not only model accuracy but also distributive and procedural justice in how robots act.

Auditability is closely related to fairness. Stakeholders need to understand why an embodied system took an action, especially when harm occurs. Action-aware memory can support auditability by preserving action histories and decision rationales, but it can also create

privacy risks if memories contain sensitive personal data. Policies must define retention limits, access rights, consent mechanisms, and redress procedures. The naming of biased commercial systems has demonstrated that public accountability can drive change, but it also requires institutions capable of interpreting technical evidence [15]. For robotics, auditability must span software, hardware, operators, and organizational protocols.

Policy frameworks for embodied autonomy are still emerging. They must address liability, safety certification, data protection, and cross-border deployment. A robot that learns continuously may not have a fixed behavior profile at certification time, complicating traditional approval processes. Regulators may require runtime monitoring, geofencing, human override, and incident reporting. Action-aware world models can assist by logging action-outcome pairs and flagging anomalies, but they cannot replace institutional oversight. The governance challenge is to design systems that are both adaptive and accountable, innovative and constrained.

8. Case Illustrations and Cross-Domain Comparisons

In manufacturing, action-aware world models could support long-horizon assembly tasks where robots must track partially completed products, adapt to defects, and coordinate with human workers. The action history is critical because the meaning of a current observation depends on which assembly steps have already occurred. A purely visual model may misclassify a component if it does not know whether a prior action succeeded. Here, memory of action outcomes supports error recovery and quality assurance. The infrastructure may be local and safety-critical, making latency and reliability more important than raw model scale.

In domestic assistance, the challenges shift toward social norms and open-ended tasks. A home robot may need to remember user preferences, past mistakes, and household routines. Action-aware memory can help it avoid repeating intrusive behaviors or unsafe actions. However, privacy and consent become central. The robot may observe sensitive activities, and its memory could be misused by operators or adversaries. Fairness concerns arise if the robot performs better for some users due to data availability or language patterns. Governance must therefore include user control, transparency, and limits on data retention.

In field robotics, such as agriculture, search and rescue, or environmental monitoring, action-aware world models must handle harsh conditions, communication delays, and long missions. The robot may need to operate for hours without human intervention, making robustness and energy efficiency paramount. Cross-domain comparison shows that no single architecture will suffice. The same principles of action-aware memory, uncertainty management, and governance apply, but their relative priority depends on domain constraints. This suggests a need for modular, configurable systems rather than monolithic models.

9. Evaluation, Benchmarking, and Open Challenges

Evaluating action-aware world models requires metrics beyond prediction accuracy. A model may predict future observations well yet fail to support useful planning. Evaluation should therefore include task success, sample efficiency, robustness to shift, calibration of uncertainty, memory retrieval quality, and computational cost. Long-horizon benchmarks must test whether agents can maintain commitments, recover from errors, and adapt to changing goals. They must also include human interaction and safety-critical scenarios, not only simulated navigation or manipulation. The complexity of such evaluation favors shared environments and reproducible protocols, but these must be designed to reflect real-world variability.

Another challenge is causal validity. Action-aware models may learn correlations between actions and outcomes that do not generalize under intervention. Since embodied agents act to test hypotheses, they need representations that support counterfactual reasoning. This is difficult in high-dimensional, partially observable settings. Combining model-based reinforcement learning with structured causal inference and probabilistic state estimation may help, but integration remains an open problem [4][6][20]. Evaluation should explicitly probe whether agents can distinguish correlation from causation when planning.

Finally, open challenges include lifelong memory management, human-robot co-adaptation, and regulatory validation. Lifelong memory requires forgetting mechanisms that preserve useful knowledge while discarding stale or harmful traces. Co-adaptation requires models that update in response to human feedback without becoming unpredictable. Regulatory validation requires evidence that systems remain safe across diverse conditions. These challenges are not only technical; they require interdisciplinary collaboration among robotics, machine learning, law, ethics, and human factors.

10. Conclusion

Action-aware world models represent a significant step toward embodied systems capable of long-horizon planning and decision-making. By treating actions as constitutive of memory, prediction, and belief revision, they address limitations of observation-centric models and support more coherent behavior under uncertainty. However, their value depends on system-level design. Architecture must balance fidelity, latency, memory, robustness, and energy. Infrastructure must support deployment, maintenance, and sustainability. Governance must ensure fairness, auditability, and accountability. The future of embodied robotics will not be determined solely by algorithmic breakthroughs, but by the institutional and technical frameworks that make action-aware autonomy safe, equitable, and maintainable. Researchers and policymakers should therefore approach action-aware world models as socio-technical infrastructures whose success is measured not only by task performance but by their contribution to trustworthy and sustainable embodied intelligence.

References

1. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
2. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.
3. Ha, D., & Schmidhuber, J. (2018). Recurrent world models facilitate policy evolution. *Advances in Neural Information Processing Systems*, 31, 2450–2462.
4. Hafner, D., Lillicrap, T., Ba, J., & Norouzi, M. (2020). Dream to control: Learning behaviors by latent imagination. *International Conference on Learning Representations*.
5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
6. Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1–2), 99–134.
7. Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe,

- D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., & Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587), 484–489.
8. Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
 9. Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, 17(39), 1–40.
 10. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
 11. Tobin, J., Fong, R., Ray, A., Schneider, J., Zaremba, W., & Abbeel, P. (2017). Domain randomization for transferring deep neural networks from simulation to the real world. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 23–30.
 12. Sadeghi, F., & Levine, S. (2017). CAD2RL: Real single-image flight without a single real image. *Robotics: Science and Systems*.
 13. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.
 14. Xiong, Zhexiong, et al. "ActWorld: From Explorable to Interactive World Model via Action-Aware Memory." *arXiv preprint arXiv:2606.17730* (2026).
 15. Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–435.
 16. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.
 17. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems*, 28, 2503–2511.
 18. Dean, J., & Barroso, L. A. (2013). The tail at scale. *Communications of the ACM*, 56(2), 74–80.
 19. Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47(1–3), 139–159.
 20. Thrun, S., Burgard, W., & Fox, D. (2005). *Probabilistic robotics*. MIT Press.