Time to Take Embodiment Seriously

John D. Martin
Department of Computing Science
University of Alberta / Amii
Edmonton, AB, Canada
jmartin8@ualberta.ca

Abstract

Reinforcement Learning (RL) systems have the potential to accomplish more than is currently possible. A major limitation, as I will argue, is their inability to leverage their bodies and the environment as pathways to agency. In what follows, I describe two different views on this issue with examples from the current literature. I show that the issue relates to the question of how the agent-environment interface affects decision making. One view believes the agent interface performs an instrumental role in the computation of decision making. Another view argues by contrast that the agent interface plays a minimal role beyond channeling sensory information and motor commands. Adopting the instrumentalist view, as I will argue, means taking embodiment seriously. Doing so has the potential to broaden the capability of RL systems, and improve RL as a model for agency.

Keywords: Embodiment, Agent-Environment Interface, Generality, Agency, Practice of RL

Acknowledgements

I would like to thank Joseph Modayil, Will Dabney, Dan Singer, Shruti Mishra, and Michael Bowling for insightful discussions that helped clarify my thoughts on these topics.

1 Introduction

Reinforcement Learning (RL) has proven to be an effective paradigm for learning to achieve goals. Still, the paradigm is limited because of a practice that fails to consider the myriad ways agents are embodied and, in addition, how an agent's body can facilitate useful forms of interaction. A reinforcement learner's body here is given by the interface the agent shares with its environment—comprised of the observation and action streams. From one view, the body provides functional utility to the processes that implement decision making. I will call this the *instrumental* view. An alternative view, found in the everyday practice of RL, adopts a comparatively orthodox stance that an agent's body contributes to decision making only by supplying information from sensors, or by executing motor commands from the policy. I will call this the *informational* view.

Arguably, what separates these views is a disagreement about how an agent's body manifests in the design of RL systems. This division can be traced back to the more fundamental question of what exactly constitutes an agent. Though I examine the connections between embodiment and agency in my article, my intention is not to adjudicate what it means to be an agent. Instead, I concentrate on describing an instrumental perspective of embodied RL, and show how this view contrasts with the everyday practice of RL. Given these two competing views of an agent, the question of which view should guide the future practice of RL naturally arises. I will not attempt to completely resolve this question here, but will point to ways the instrumental vision could guide the design of new, general-purpose RL algorithms. Ultimately, what is at stake is a consensus about the fundamental nature of an agent, and how the community should go about designing its RL systems that scale and generalize to increasingly complex environments. By pointing out that embodiment has important ramifications on agency, I hope to convey that the topic deserves our serious attention and continued research.

2 Agency, Embodiment, and Reinforcement Learning

Agency can be understood from the interaction between a system and its environment. An agent is a system that receives observations and returns actions. Natural agents like humans take in vast amounts of observations in the form of light intensity, changes in air pressure, deflections of hair follicles, fluctuations in skin temperature, the stimulation of taste and olfactory receptors; they return actions through muscle twitches and other bodily functions. Of course not everything is an "agent." For instance, inert objects like rocks, substances like water, and plumes of smoke are not agents. These things have no agency, because they have no potential to sense the world and accomplish anything by themselves (Minsky, 1988).

The connection between embodiment and the potential do something (i.e. agency) has deep roots in the cognitive and computational sciences, but also in the teachings of Buddhist meditation, and in the philosophies of Hursserl, Merleau-Ponty, Varela et al. (1991), Clark (1997), and Kiverstein (2012). Embodied agents posses a body with which they can affect, and be affected by their environment. The body's unique coupling to an environment affords an agent with different ways of interacting that ultimately determine the agent's decision making potential. For example, the interactions afforded to a manipulator robot anchored to a factory floor are much different than a mobile manipulator, which can move to different locations. In some sense, the mobile robot has more agency, because it has the potential to accomplish more than one anchored to a single location. This example also illustrates how agency has range; it is not a simple binary property.

In reinforcement learning, an agent's body appears in two parts of the standard formalism. The first is in the agent's joint transition dynamics with the environment—things like mass, the material properties of its construction, and how the observations evolve in response to actions (Pfeifer and Bongard, 2006). The second place the body appears is in the construction of the agent-environment interface. The interface is defined by the observation stream that contains incoming sensory data, and the action stream that outputs information to motors. Together these constitute an agent's *stream of experience*. Ways in which an RL agent's body contributes to decision making will be made clear in the next section.

3 Ways the Body Contributes to Decision Making

In H.G. Wells' science fiction novel, *The Island of Doctor Moreau*, a mad scientist creates hybrid animals through the practice of vivisection—surgically combining parts of different animals to create a new species. Although such creations are biologically implausible, not to mention disturbing to consider, Wells underscores several important ways in which the body participates in decision making. The first example highlights how an organism's abilities are dependent on its body's shape and substance—its morphology. This mirrors work within robotics on under-actuated systems (Tedrake, 2009) which shows how exploiting a body's natural dynamics can produce a useful form of action. Another aspect of embodied agency in Wells's creatures is how his organisms' subjective experience comes from their uniquely-structured interface with the environment. Evidence of this connection can be found in several works from the field of cognitive neuroscience, e.g. Olshausen and Field (1996). The final example Wells provides is more commonplace; it is the idea that a body can co-opt technology and tools from the environment to transform its decision making capabilities. This point has been emphasized in the work of animal tool use (Seed and Byrne, 2010) and Extended Intelligence. In what follows, I examine how each of these examples contributes to decision making and how they appear in the everyday practice of RL.

3.1 Morphology

The first question I consider is how can an agent's bodily shape, material properties, and sensorimotor typology—the things that define its morphology—contribute to its decision making potential. The body participates through two distinct modes of action. One, which I will call *computational action*, is commanded from the learning system's policy to override the body's natural responses—something control theorists call "high gain commands" (Tedrake, 2009). Here we actively overpower the system to achieve a desired outcome. Alternatively, an embodied agent can passively act through its body's natural dynamics. These *morphological actions* can be executed even when the learning system is turned off. Although the consequences of each action may appear equivalent, in many cases, morphological action can improve efficiency and amplify the agent's natural abilities (Pfeifer et al., 2014).

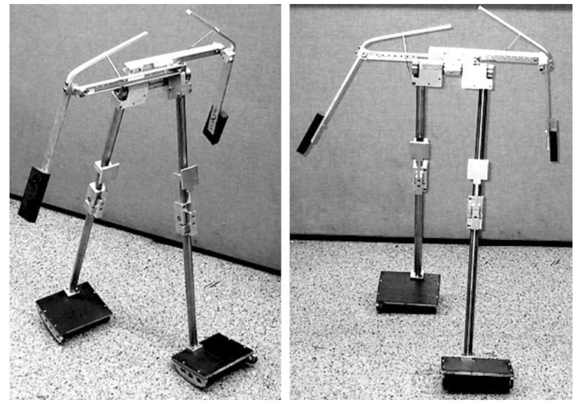

Figure 1: Passive walkers (Collins et al., 2001) illustrate how body **morphology** can implement decisions otherwise imposed with computation.

Support for this claim can be found from work within robotics, observations of nature, and regular human activity. Passive dynamic walkers, shown in Figure 1, are mechanisms that achieve stable bipedal motion through strict morphological action—without the use of sensors or actuators—only by the shape and arrangement of their bodies (Collins et al., 2001). Bats, for instance, achieve incredible maneuverability by tapping into the potential afforded by their body mechanics; they can pivot 180-degrees while flying at full speed using only a small amount of space and actuation (Tian et al., 2006). Birds also make efficient use of their bodies and environment while gliding. Finally, consider a technique known by many rock climbers: that grasping holds requires much less effort if their skeletal structure can take weight away from their muscles. These examples illustrate how morphological actions can support proficient decision making.

The everyday practice of RL provides examples of how morphological action is viewed as an aspect of the environment, rather than of the agent. In the canonical Mountain Car domain, for instance, an agent uses momentum to amplify the effect of its computational actions (horizontal acceleration) (Sutton and Barto, 2018). In this case, morphological actions are considered part of the environment—in line with the informational view that believes action only occurs when the learning system commands it. Furthermore, Mountain Car has been constructed so a morphological affordance is always available. This leaves open the question of how RL systems can and should be constructed to efficiently exploit their body's natural dynamics. In principle, morphological action can be realized through direct trial-and-error learning—provided it is aligned with the preferences of the reward signal. But does all morphological action emerge this way? Or are there general-purpose algorithms that can redraw the agent-environment interface to include the actions the body induces?

3.2 Structure

Bodies also equip agents to act in specific situations by structuring information from their stream of experience. At a basic level, structure is imposed by the very arrangement of sensors and motors. This can have the effect of correlating nearby observation signals and motor outputs within spatial settings. Generally speaking, a body's distribution of sensors, by virtue of its morphology and the agent's goals, imposes the relationships that structure experience. Structuring is a part of how an agent detects affordances or the features of a situation to ultimately inform action. Furthermore, the body fluently re-structures the experience stream in response to a dynamically changing environment.

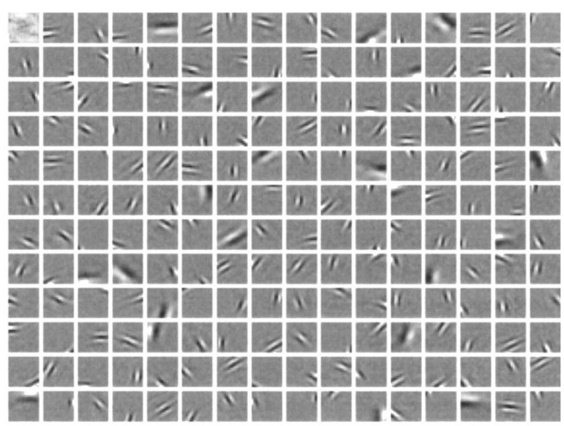

Figure 2: Receptive fields (Olshausen and Field, 1996) structure **the observation stream**, similar to that found in the primary visual cortex.

We can find some evidence of these claims in the cognitive neuroscience literature on vision. For instance, the region of the brain that governs visual saccades—the rapid positioning of an eye's fovea to new target locations—can be controlled by multiple sensory modalities. Humans often saccade to the origin of sounds, things detected only by touch, or visual cues from peripheral vision (Clark, 1997). Not only are saccades controlled by multiple modalities, but target locations are based on the subsets of sensory stimuli structured by the body. Receptor fields (subsets of raw sensory input) are another way bodies structure observational information. Olshausen and Field (1996) showed in their seminal work that maximising sparseness of neural activity can lead to the emergence of localized receptive fields, which are spatially structured—like those found in the primary visual cortex (Figure 2).

How is the structure of experience treated in practice? Most often, the observation and action streams are viewed as fixed elements of the environment that a human specifies and provides to the agent. This practice aligns with the informational view, that the body exists as a passive channel of sensorimotor information. A clear example can be found from computer science, particularly in deep RL, where it is often assumed the agent senses the world through a constant stream of camera imagery. Since pixel relationships are evident by virtue of an image's construction, deep RL agents can be designed with inductively-biased sparse architectures, whose generalization properties are advantageous for sample-efficient learning (Mnih et al., 2015). The view implicitly expressed here is that structuring experience is the responsibility of the human—to encode then make available through the environment interface. Such biases ultimately limit agency, because the agent can only apply to situations where its structural assumptions hold. Furthermore, there are situations where a structural bias will be unavailable or prohibitively difficult to encode, such as in learning visual saccades from multiple sensory modalities. In these situations the community has few methods to use.

Current work on making sense of sensory data can be found within computer science, specifically neural architecture search (Hoefler et al., 2021). In RL, Martin and Modayil (2021) demonstrate a prediction algorithm that is able to impose its own prediction-based observation structure through a process of auxiliary prediction. There are additional situations to consider, such as when observational structure is available, but common biases may not be the most useful. Also in the long-life view of decision making, observational structure will surely change. For instance, a robot may be upgraded multiple times in its life while encountering failures and other changes to its stream of experience. All these examples underscore ample opportunity for future research on body-agnostic algorithms that can apply to any experience stream.

3.3 Extension

I also claimed that agents can transform their bodies by co-opting external resources, such as tools and technologies, into their decision making. This claim aligns with in the Extended Mind hypothesis put forth by Clark and Chalmers (1998). As agents include tools and even other agents as part of their action model, they modify their affordances. When these are properly exploited, agents can accomplish more than they otherwise could.

We do not have to look far to find examples of this principle. Humans often use notebooks to externalize their thoughts and retrieve them at a later time. Chimpanzees and capuchin monkeys use stone tools to both dig and crack nuts open (McGrew, 2010). Caledonian crows provide another remarkable example of animal tool use (Figure 3).

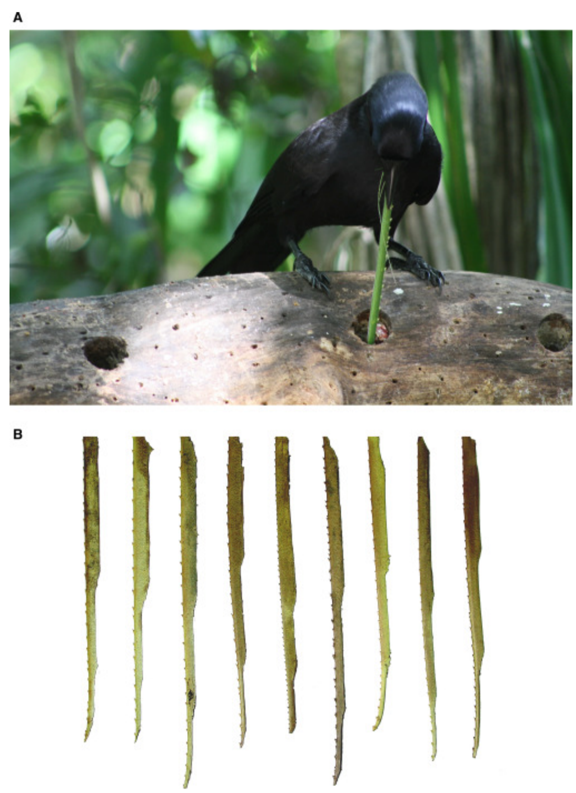

Figure 3: Caledonian crows (A) use tools (B) to **extend** their capabilities (Seed and Byrne, 2010).

How do we place the Extended Mind hypothesis in relation to the daily practice of RL? Previous work in computer science is an expression of the informational view, which demonstrates how RL agents can learn to use tools when it explicitly aligns with their reward (Allen et al., 2019), or when tools are specified in the agent-environment interface (Baker et al., 2019). Future work in this area remains open.

4 Final Reflection

So should the daily practice of RL be guided by an instrumental view of embodied agency, or is it business as usual? Upon examining the ways in which embodiment matters for decision making, we observed new benefits that could improve the design of RL systems, and the adequacy of RL as a general model for agency. Specifically, we observed that leveraging a body's natural dynamics brings additional efficiency and capability through morphological action; structuring the stream of experience provides an agent with generality to achieve goals in host of different settings; and that co-opting tools from the environment can transform and simplify problems. If the goal is to understand ways to design general-purpose RL systems with body-agnostic algorithms, then the instrumental view seems germane. Though the need for generality in the design of embodied learning systems remains hotly debated (Kaelbling, 2020; Roy et al., 2021), one thing seems clear; embodiment has important ramifications on agency, and the topic deserves our serious attention and continued research.

References

Allen, K. R., Smith, K. A., and Tenenbaum, J. B. (2019). The tools challenge: Rapid trial-and-error learning in physical problem solving. *arXiv preprint arXiv:1907.09620*.

Baker, B., Kanitscheider, I., Markov, T., Wu, Y., Powell, G., McGrew, B., and Mordatch, I. (2019). Emergent tool use from multi-agent autocurricula. *arXiv preprint arXiv:1909.07528*.

Clark, A. (1997). *Being there*. MIT Press Cambridge, MA.

Clark, A. and Chalmers, D. (1998). The extended mind. *analysis*, 58(1):7–19.

Collins, S. H., Wisse, M., and Ruina, A. (2001). A three-dimensional passive-dynamic walking robot with two legs and knees. *The International Journal of Robotics Research*, 20(7):607–615.

Hoefler, T., Alistarh, D., Ben-Nun, T., Dryden, N., and Peste, A. (2021). Sparsity in deep learning: Pruning and growth for efficient inference and training in neural networks. *Journal of Machine Learning Research*, 22(241):1–124.

Kaelbling, L. P. (2020). The foundation of efficient robot learning. *Science*, 369(6506):915–916.

Kiverstein, J. (2012). The meaning of embodiment. *Topics in cognitive science*, 4(4):740–758.

Martin, J. D. and Modayil, J. (2021). Adapting the function approximation architecture in online reinforcement learning. *arXiv preprint arXiv:2106.09776*.

McGrew, W. C. (2010). Chimpanzee technology. *Science*, 328(5978):579–580.

Minsky, M. (1988). *Society of mind*. Simon and Schuster.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533.

Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.

Pfeifer, R. and Bongard, J. (2006). *How the body shapes the way we think: a new view of intelligence*. MIT press.

Pfeifer, R., Iida, F., and Lungarella, M. (2014). Cognition from the bottom up: on biological inspiration, body morphology, and soft materials. *Trends in cognitive sciences*, 18(8):404–413.

Roy, N., Posner, I., Barfoot, T., Beaudoin, P., Bengio, Y., Bohg, J., Brock, O., Depatie, I., Fox, D., Koditschek, D., et al. (2021). From machine learning to robotics: Challenges and opportunities for embodied intelligence. *arXiv preprint arXiv:2110.15245*.

Seed, A. and Byrne, R. (2010). Animal tool-use. *Current biology*, 20(23):R1032–R1039.

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

Tedrake, R. (2009). Underactuated robotics: Learning, planning, and control for efficient and agile machines course notes for mit 6.832. *Working draft edition*, 3.

Tian, X., Iriarte-Diaz, J., Middleton, K., Galvao, R., Israeli, E., Roemer, A., Sullivan, A., Song, A., Swartz, S., and Breuer, K. (2006). Direct measurements of the kinematics and dynamics of bat flight. *Bioinspiration & biomimetics*, 1(4):S10.

Varela, F. J., Thompson, E., and Rosch, E. (1991). The embodied mind: Cognitive science and human experience.