A Braitenberg vehicle's habitat

Manuel Baltieri[1,2,*], Fernando E. Rosas[2−5], and Filippo Torresan[1,*]
[1]Araya Inc., Japan
[2]Department of Informatics, University of Sussex
[3]Sussex AI and Sussex Centre for Consciousness Science, University of Sussex
[4] Department of Brain Sciences, Imperial College London
[5]Centre for Eudaimonia and Human Flourishing, University of Oxford

Corresponding authors: manuel_baltieri@araya.org, filippot.research@protonmail.com

Abstract

Agents do not generally interact with the entirety of their physical environment equally. What matters to them depends on their morphology, sensory interface, embodied dynamics, and action repertoire. This paper formalises some of these ideas and illustrates them for Braitenberg vehicles. We distinguish the agent's physical environment from its habitat, understood as the effective world disclosed once particular constraints of a class of agents are taken into account. We then provide a formalisation of physical environment and habitat, and describe a systematic way to construct the latter from the former. To do so, we model the agent's environment as a partially observable Markov decision process (POMDP) and describe the habitat as another POMDP in the form of a quotient or coarse-grained version of the environment. This coarse-graining is obtained by identifying environmental differences that induce the same observations and coarse-grained dynamics under every admissible action. For the standard 2b vehicle in an isotropic light field, the resulting habitat removes absolute position around the source while retaining distance and relative heading. We then sketch three stylised variations of vehicle 2b, with analogies to a sunflower, a sea urchin, and a crab, to illustrate how different morphologies, sensory interfaces, dynamics, and action structures suggest different candidate habitats. This provides a possible operationalisation of a habitat for minimal agents, under an explicit modelling choice about the agent-environment boundary.

Code available at: https://github.com/mbaltieri/BraitenbergHabitat

Introduction

The notion of an environment *for* an agent, i.e. perceptual features and actions constructed or carved out from its larger surroundings based on the agent's sensory and motor apparatuses, is at the centre of various treatments of embodied (Clark, 1997; Baggs and Chemero, 2021) and enactive (Varela and Rosch, 1992; Thompson, 2007; Di Paolo et al., 2017) cognitive science, as well as of earlier work that influenced those traditions, see e.g. Merleau-Ponty (1942). This is often captured by the concept of *umwelt*, describing a closed unit of perception and action that characterises an agent as it navigates its surroundings (von Uexküll and Kriszat, 1934; von Uexküll, 2010).

The term "umwelt", however, has been used to refer to different ideas over the years, perhaps even by von Uexküll himself (Brentari, 2013, 2015; Feiten, 2020, 2024). One of these interpretations is more ontological, relating "umwelt" to some notion of subsystem of the physical environment that is relevant to a typical agent (e.g. a typical living organism belonging to a species), and is closely related to affordances in ecological psychology (Gibson, 1979). Another interpretation leans towards an epistemic stance, aiming to characterise an agent-centric perspective of what may constitute an effective environment that is relevant from an agent's subjective point of view, and better aligns with a constructivist position typical of enactivist approaches to sense-making (Varela and Rosch, 1992; Thompson, 2007; Di Paolo et al., 2017). In an attempt to reconcile the insights of ecological psychology and enactivism, recent work has argued that umwelt should be distinguished from both i) the environment as physical reality (here simply addressed as *environment*) and ii) the environment as *habitat* for an agent (Baggs and Chemero, 2019, 2021), see also (Feiten, 2020, 2024) for a similar distinction. On this view, a physical reality corresponds to everything that exists across different spatial and temporal scales of a particular choice of universe (e.g. the one we inhabit, or an artificial one specified by the rules of a cellular automaton). A habitat, on the other hand, is a part of that reality specified by the constraints of how some groups of agents operate, e.g. the members of an animal species that can perceive and make use of affordances at a certain spatio-temporal scale. Following this, the umwelt would then be a part of the habitat, the environment *for* an agent shaped by its individual learning and sensorimotor history. Habitat and umwelt thus play importantly different roles. An umwelt is the individually enacted perception–action world of an organism and may therefore vary with its action policies, sensorimotor history, and specific circumstances. A habitat is less individualised, in the sense that it is the part or scale of physical reality that defines the effective physics of a species or agent type, given constraints common to agents of the same type, such as its morphology, sensory interface, and action repertoire.

These conceptual proposals do not, however, clarify in what sense the habitat is part of the environment and how the former is carved out from the latter. In this work, we formalise this by taking a common definition of environment as a partially observable Markov decision process (POMDP), and introducing habitat as another POMDP obtained as a coarse-graining of the environment POMDP. This coarse-graining aggregates states that are indistinguishable from the perspective of a class of agents, in the sense that they lead to the same observations for agents of the same kind. Notably, in this formalisation a habitat is not a subsystem (a subset of states and actions closed under transition and observation maps) or a component (a system obtained from a factorisation) of the environment, as it is sometimes understood in the literature. It is a quotient that retains the environmental distinctions and dynamics available at the level of a relevant agent–environment interface.

Our exposition focuses on minimal models of autonomous agents, Braitenberg vehicles (Braitenberg, 1984), with the aim of studying in detail examples of structural constraints on how these simple agents produce actions from sensory readings or observations. By doing so, we shed light on the question of how the specification of such constraints affects a vehicle's habitat, while leaving a formalisation of umwelt to future work.

Related work

Our investigation of habitats is inspired, on the technical side, by the idea of bisimulation equivalence in concurrent systems (Sangiorgi, 2009), and particularly by its categorical generalisations (Rutten, 2000b; Jacobs, 2017) with applications in reinforcement learning, to Markov decision processes (MDPs) (Zhang et al., 2021a,b) and partially observable MDPs (POMDPs) (Castro et al., 2009; Rosas et al., 2025; Zhang et al., 2026), and neuroscience (Baltieri et al., 2025b).

Studies on observers in control theory (Wonham, 1976; Hepburn and Wonham, 1984) propose a similar approach that predates much of the work on bisimulation equivalences as discussed here, but the connection is only made clear using generalisations that came later, see for instance (Rutten, 2000b; Baltieri et al., 2025a). Informally, bisimulation equivalences capture the idea that a given (dynamical) system can be quotiented by an equivalence relation on the states of the system that generate outputs that are not distinguishable by an observer or agent, while preserving the system's transition and output maps. The notion of *indistinguishability* depends on the choice of the equivalence relation and on what is considered to be the output of the system.

Building on proposals in reinforcement learning and neuroscience, in this work we use these ideas to formalise the fact that an agent can only distinguish certain aspects of its environment while interacting with it. Consider for instance the notion of visible spectrum, the band of the electromagnetic spectrum that is visible to the human eye. This already captures a definition of indistinguishability for the human eye by specifying parts of the spectrum that are outside the visible one. This definition is not tied directly to the environment, but to the way an agent interacts with it. Consider for instance how some insects we share an environment with have a rather limited visual understanding of our visible spectrum, but at the same time possess the ability to see parts of the ultraviolet spectrum which are not directly visible to us (Philippides et al., 2011; Cronin and Bok, 2016). Building on this idea, we assume, for now informally, that if a change in the environment is registered by an agent, in the sense that it can be observed or measured by it, then this can be used to describe part of the environment *for* an agent.

This work is also inspired by Ay and Löhr (2015), who formalise a definition of umwelt in a measure-theoretic framework, focusing on the difference between extrinsic and intrinsic perspectives. We believe that our definition of habitat is closely related to a non-measure-theoretic version of their notion of extrinsic perspective.

Technical background

We now review a few core parts of standard reinforcement learning setups, which constitute the backbone of our technical proposal. We start with an environment, presented as a POMDP (Puterman, 2014; Kaelbling et al., 1998; Sutton and Barto, 2018). A POMDP is a way of describing a stochastic open dynamical system in terms of stochastic hidden states, inputs, and outputs. Here, we use this as a description of the physical ***environment*** of the agent.

Definition 1 (Environment as a POMDP). A partially observable Markov decision process is a tuple $\mathcal{Q} = (S, A, T, O, M)$, where:

- S is the state space,
- A is the action space,
- O is the observation space,
- $T : S \times A \to P(S)$ is the transitions map, where $P(S)$ is the set of distributions over S such that for a given state s and action a, $T(s, a)$ gives a probability distribution of states an agent can transition to from state s while taking action a, often written as $p(s' \mid s, a)$,
- $M : S \to P(O)$ is the observation map where $P(O)$ is the set of distributions over O such that for a given state s, $M(s)$ gives a probability distribution on observations $p(o)$ for a given state $s \in S$, $p(o \mid s)$.

In standard (online) reinforcement learning, an agent uses its sensorimotor history, given by sequences of past observation-action pairs, to achieve its goals by learning a policy often expressed as a map $\pi : B \to P(A)$, where B is the space of belief states, sufficient statistics (in a Bayesian sense) of histories $h_t := (o_0, a_0, \dots, o_{t-1}, a_{t-1}, o_t)$ of action-observation pairs (Kaelbling et al., 1998; Murphy, 2025). The policy π is selected to maximise the expected cumulative discounted reward, also known as expected return, based on past and current experiences acquired through meaningful interactions with the environment. [1]

In practice, however, an agent often only interacts with parts of its environment. This could be due to spatial constraints (localised interactions), time constraints (different time scales between parts of the environment and an agent) or simply to having different interfaces (limiting what the agent can observe (Rosas et al., 2025)). This implies that much of the structure of a particular environment may be irrelevant for an agent, at least for a specific task (Zhang et al., 2021b; Javed and Sutton, 2024). To formalise this idea as a form of agent-relative environmental abstraction, it is useful to consider the notion of bisimulation equivalence on (PO)MDPs (Givan et al., 2003; Castro et al., 2009; Rosas et al., 2025).

Definition 2 (**Indistinguishable states as a bisimulation equivalence**)**.** Let $\mathcal{Q}$ be a POMDP and $B \subseteq S \times S$ be an equivalence relation of states of $\mathcal{Q}$. An equivalence relation B is a bisimulation equivalence on $\mathcal{Q}$ (Rosas et al., 2025) if, whenever two states $s, s' \in S$ are equivalent, $(s, s') \in B$, the following conditions are satisfied for $p(\cdot) \in P(\cdot)$:

$$\begin{aligned} &(i) \;\; p(o \mid s) = p(o \mid s') && \forall o \in O \\ &(ii) \;\; p(c \mid s, a) = p(c \mid s', a) && \forall c \in S/B, \forall a \in A \end{aligned}$$

where S/B is the set of equivalence classes and $p(c \mid s, a) := T(s, a)(c)$ denotes the mass assigned by the transition kernel $T(s, a)$ to the equivalence class $c \in S/B$. In the discrete case, this reduces to $p(c \mid s, a) = \sum_{\bar{s} \in c} p(\bar{s} \mid s, a)$.

Informally, two states are equivalent in this sense if they look the same to the agent and, under any fixed action, evolve in the same way at the coarser level defined by the equivalence classes. Interestingly, for the fixed reactive policy considered below, the first condition on equal observations also implies equal realised actions. The second condition, on the other hand, quantifies over all possible actions $a \in A$, regardless of whether they are realised by a particular agent.

Remark 3 (Different notions of bisimulation equivalence)**.** One might weaken condition (ii) by quantifying only over actions associated with a given fixed policy π. We do not adopt this alternative for three reasons. First, tying the equivalence relation to a single fixed policy would make the definition difficult to apply when the policy changes through learning (Sutton and Barto, 2018), or when an agent can select among several policies (Levine et al., 2020). This problem does not arise for the Braitenberg vehicles considered here, which have fixed deterministic policies, but we want the proposed notion of habitat to apply beyond this special case. Second, if only the actions actually realised from a particular initial condition were considered, the resulting equivalence relation would depend on the trajectory generated from that condition. Even under the same policy, different initial states may lead to different observations and realised actions, and hence to different quotients. A habitat would then vary between different runs of the same agent, or even between different initial conditions for the same individual, rather than describing an environment relevant to a class of agents. Such dependence on individual history is more naturally associated with umwelt. Third, one might instead quantify over the complete range or support of π. This avoids dependence on a single trajectory, but it still considers, at a given state and time, actions that the policy selects only under other observations or histories and therefore could not select in the present circumstances. [2] We therefore define habitat at the level of the embodied observation and action interface, and ask which environmental distinctions remain irrelevant under every action admitted by that interface. Policy- and history-dependent reductions may instead contribute to a future formalisation of umwelt.

Definition 2 is given at the level of general stochastic POMDPs, where bisimulation is expressed by equality of observation laws and equality of transition mass assigned to equivalence classes. For the Braitenberg vehicles considered below, this level of generality is not needed, since their dynamics and observation equations are deterministic. In that setting, the bisimulation conditions can be rewritten more concretely in terms of equality of observations and equality of quotient successors. The following remark records this deterministic form of Definition 2, which is the form used in the subsequent analysis of vehicle 2b.

Remark 4 (Deterministic specialisation)**.** In the deterministic case considered below for Braitenberg vehicles, there exist maps $F : S \times A \to S$ and $G : S \to O$ such that $T(s, a) = \delta_{F(s,a)}$ and $M(s) = \delta_{G(s)}$. Hence condition (i) reduces to $G(s) = G(s')$. Moreover, for any equivalence

[1] Since we won't be considering learning in this work, we won't spell out a formal definition of expected return, and similarly we won't consider rewards and discount factors directly. It is however not difficult to think of a POMDP's observations as a combination of a quantity observable by the agent and a reward signal, so that $O = O' \times$ Reward, meaning that rewards can in general be seen as folded in the observation space.

[2] A possible way to handle this is to make the available actions depend on the current state (Bakirtzis et al., 2025). On this account, MDPs are formulated using a state–action space A equipped with a map $\psi : A \to S$, so that the actions available at $s \in S$ form the fibre $A_s = \psi^{-1}(s)$, rather than assuming that the same action set is available at every state. An analogous extension to POMDPs could allow a bisimulation relation to compare only the actions available at each particular state. We leave this extension to future work.

class $c \in S/B$, we have

$$p(c \mid s, a) = \delta_{F(s,a)}(c) = \begin{cases} 1 & \text{if } F(s,a) \in c, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore condition (ii) reduces to the requirement that, for every $a \in A$, the successors of s and s' lie in the same equivalence class. Equivalently, if $q : S \to S/B$ is the quotient map, then condition (ii) can be written as

$$q(F(s,a)) = q(F(s',a)) \qquad \forall a \in A.$$

As shown, for instance by Rutten (2000a), a bisimulation equivalence induces a quotient POMDP whose states are the corresponding equivalence classes, with transition and observation maps inherited from the original POMDP. Quotient POMDPs can thus be used to describe the *effective world* for an agent solving a particular problem. This will give us a formal notion of ***habitat***, where the relevance of the environment *for* the agent is captured by a bisimulation equivalence, i.e. where irrelevant information is quotiented out of a given POMDP.

Definition 5 (**Habitat as a quotient POMDP**)**.** Let $\mathcal{Q}$ be a POMDP and let B be a bisimulation equivalence on $\mathcal{Q}$. The induced quotient POMDP is a POMDP Q' with state space $S' := S/B$, the same action and observation spaces, transitions $T' : S' \times A \to P(S')$ where $p(c \mid [s], a) := p(c \mid s, a)$ with $[s]$ is the equivalence class s belongs to, and observation map $M' : S' \to P(O)$ where $p(o \mid [s]) := p(o \mid s)$.

These assignments are well-defined because, by Definition 2, equivalent states induce the same observation law and the same transition law on equivalence classes. One can think of a quotient POMDP as a POMDP whose state space is a (state) coarse-graining of a given POMDP, i.e. a structure-preserving coarse-graining (a surjective map on states) of the original POMDP, generalising the equivalent definition for MDPs (Ravindran and Barto, 2003). Importantly, different bisimulation equivalences on the same environment POMDP may induce different quotient POMDPs. We therefore do not assume that a habitat model is unique or that a chosen quotient is maximally coarse. We also note that our definition makes some simplifying assumptions, with action and observation spaces equal to those of the original POMDP, cf. (Ni et al., 2024).

Remark 6 (Different notions of bisimulation equivalence)**.** An on-policy bisimulation (Castro, 2020) offers another possibility: it compares only the transitions produced when a fixed policy is followed. Besides the difficulties discussed in Remark 3, however, this changes the kind of quotient produced. Once the policy π has been incorporated into the transition law, and in the stochastic case its actions have been averaged out, the resulting system no longer accepts actions from the original action space A. Its natural quotient is therefore an autonomous Markov process with observations, rather than a POMDP retaining the environment's original action interface. Formally, this process can still be represented as a POMDP with the singleton action space $A = \{*\}$. Such a representation is degenerate, however, because the alternative actions that could have been applied to the original environment are no longer present. Our aim is instead for a habitat to remain the same general kind of object as the environment, i.e. a POMDP that takes a current state and an action and determines distributions over the next state and observation. This allows the habitat to be interpreted as a restricted version of the environment, rather than merely as the process generated by one particular policy. Other choices are possible, but they are not considered in the present construction.

This notion of habitat can thus be interpreted here as a coarse-grained version or abstraction of the environment in which the agent acts. More precisely, the quotient POMDP should be read here as a formal model of habitat for a given environment, i.e. relative to a specified agent-environment factorisation or boundary, rather than as a stand-alone notion of "habitat". The coarse-graining that shapes the habitat is relative to the embodied interface encoded by the POMDP: its morphology, sensor modalities, observation map, transition dynamics, and action space, as we explore next.

A Braitenberg vehicle's environment

POMDPs are popular in model-based reinforcement learning (Sutton and Barto, 2018) and cognitive neuroscience (Parr et al., 2022) as a way to describe how agents learn a model about the environment they mainly use for planning. This is, however, seemingly at odds with some areas of artificial life research, condensing ideas from embodied and enactive cognitive science into a research program for seemingly "representation-free" agents (Harvey, 1992; Beer, 1995; Van Gelder, 1995; Beer, 2000; Harvey, 2008). A prototypical example of this program is given by Braitenberg vehicles (Braitenberg, 1984), agents that solve tasks involving running towards or away from a particular stimulus using a rather trivial, stateless, configuration of a few wires that appear to suggest no internal model is employed or even needed to display adaptive and goal-directed behaviour. For a recent review of their applications see (Shaikh and Rañó, 2020), and for analyses of various technical aspects of these vehicles' behaviour, see for instance (Rañó, 2012; Hotton and Yoshimi, 2024).

In what follows, we will use POMDPs as part of a mathematical formalism to describe environment and habitats of some standard Braitenberg vehicles, and some new variations we introduce here. This should not be interpreted as suggesting that these simple agents learn or store some kind of internal model of their environment or habitat. Furthermore and perhaps counter-intuitively, while using tools from reinforcement *learning*, we will not in fact consider any standard notion of learning in our treatment. Our models of environment and habitat ought to be simply regarded as a (convenient in our opinion) way of specifying the constraints on the sensorimotor couplings of these agents.

Table 1: Symbols used in the environment and habitat models.

Symbol	Meaning
$X_{\mathsf{Centre}} \subseteq \mathbb{R}^2$	Vehicle centre position
$X_l, X_r \subseteq \mathbb{R}^2$	Left/right sensor positions
$\Theta \subseteq (-\pi, \pi]$	Absolute heading
$\Theta^{\mathsf{rel}} \subseteq (-\pi, \pi]$	Signed heading relative to the stimulus
$X_{\mathsf{Stim}} \subseteq \mathbb{R}^2$	Stimulus position
$R_{\mathsf{Body}} \subseteq \mathbb{R}_{>0}$	Body radius
$V_l, V_r \subseteq \mathbb{R}$	Left/right wheel speeds
$O_l, O_r \subseteq \mathbb{R}$	Left/right sensor readings
$R \subseteq \mathbb{R}_{>0}$	Source-centred radius of X_{Centre}
$R_l, R_r \subseteq \mathbb{R}_{\geq 0}$	Source-centred radii of the sensors
$\Phi \subseteq (0, 2\pi]$	Source-centred polar angle of X_{Centre}

Braitenberg vehicles 2 and 3 are classical examples of minimal adaptive agents that include only two sensors (e.g. stimulus sensors), one mounted on each side, and two motors, each driving one of the left or right wheels. Their behaviour is enacted by fixed connections between sensors and motors, which may be excitatory or inhibitory and either ipsilateral (same-side) or contralateral (cross-side). Their environment can be described as a deterministic POMDP, given next. [3]

For this, fix once and for all a stimulus position $x_{\mathsf{Stim}} \in X_{\mathsf{Stim}}$ and a body radius $r_{\mathsf{Body}} \in R_{\mathsf{Body}}$. We assume $x_{\mathsf{Stim}} \notin X_{\mathsf{Centre}}$ and that the transition dynamics preserve X_{Centre}. We also assume that X_{Centre} is invariant under rotations about x_{Stim}. These will be treated as model parameters rather than dynamical state variables. Accordingly, the models below are to be read relative to the fixed choice of parameters $(x_{\mathsf{Stim}}, r_{\mathsf{Body}})$. Maps that depend on these quantities are therefore written as parameter-indexed families, rather than as maps with X_{Stim} or R_{Body} as additional varying inputs.

Example 7 (A Braitenberg vehicle's environment)**.** Following the POMDP structure of Definition 1, we have:

- the state space $S \coloneqq X_{\mathsf{Centre}} \times \Theta$,
- the action space $A \coloneqq V_l \times V_r$,
- the observation space $O \coloneqq O_l \times O_r$,
- transition dynamics given by the product of the following maps, updating position and heading of the vehicle respectively,

$$f_{X_{\mathsf{Centre}}} : X_{\mathsf{Centre}} \times \Theta \times V_l \times V_r \to X_{\mathsf{Centre}}$$
$$f_{\Theta}^{r_{\mathsf{Body}}} : \Theta \times V_l \times V_r \to \Theta, \tag{1}$$

- an observation map given by the sequential composition of the following two maps, the first one converts the vehicle's pose (centre position and heading) to positions of the sensors and the second one gives stimulus readings at the positions of the sensors,

$$f_{\mathsf{Sensor}}^{r_{\mathsf{Body}}} : X_{\mathsf{Centre}} \times \Theta \to X_l \times X_r$$
$$f_{\mathsf{Stim}}^{x_{\mathsf{Stim}}} : X_l \times X_r \to O_l \times O_r. \tag{2}$$

This example fixes a particular factorisation of the overall system into agent and environment. The POMDP includes the vehicle's pose, body geometry, sensor geometry, and kinematics in the environment model, while its observation and action spaces specify the interface through which a controller interacts with that model. A particular Braitenberg wiring is represented by a reactive deterministic policy $\pi : O \to A$. Since the habitat relation used here quantifies over every $a \in A$, the resulting quotient is relative to the chosen interface but not to, e.g. the particular wiring of vehicle 2b. Other agent–environment factorisations are possible; for example, the body and its kinematics could instead be included on the agent side of the boundary. We leave such alternatives to future work.

The specific transition and observation maps depend on the vehicle's body and sensor geometry and on the stimulus dynamics. We model a circular differential-drive body with two lateral sensors and two wheels, assume instantaneous sensing, and omit walls, occlusion, drag, and friction. The stimulus is an isotropic point light whose intensity depends only on source distance. Symbols are listed in Table 1 while numerical parameters and implementation details are provided in the code. We now focus on vehicle 2b.

A Braitenberg vehicle's habitat

Braitenberg vehicle 2b is characterised by two excitatory contralateral connections. The behaviour of the vehicle is specified by the following policy, for $(o_l, o_r) \in O$ and $(a_l, a_r) \in A$, for some constant gain $k \in \mathbb{R}_{>0}$:

$$a_r = k \cdot o_l, \quad a_l = k \cdot o_r. \tag{3}$$

A sample trajectory is shown in Fig. 1.

This vehicle is aptly characterised as aggressive, since if the stimulus is directly in front of it both sensors receive similar inputs, both motors spin at similar rates and so the vehicle accelerates towards the stimulus, in some sense attacking it, with behaviour that appears as aggressive or attraction-seeking. If the stimulus is to the left, the left sensor receives more input, the right motor spins faster than the left and the

[3] Our definition is for discrete-time systems. We note that this isn't how Braitenberg vehicles are usually depicted, however there isn't much of a difference as long as we assume that sensory readings are *instantaneously* passed to the wheels (delay can in fact severely affect the behaviour of the vehicles).

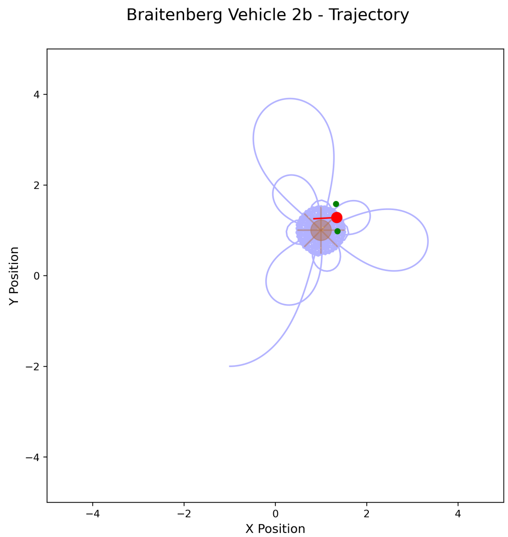

Figure 1: An example trajectory of Braitenberg's vehicle 2b.

vehicle turns toward the stimulus, vice versa if it is on the right.

The environment used by vehicle 2b is the POMDP in Example 7 with fixed parameters and maps. Its habitat, however, depends on which distinctions are relevant to this vehicle. We therefore begin by identifying states of the environment that are indistinguishable for vehicle 2b.

Remark 8 (Indistinguishable positions)**.** If the vehicle is rotated around the light source while keeping fixed its heading relative to the source, then its sensory readings do not change. This means that it cannot distinguish the corresponding change in its Cartesian position, see Fig. 2.

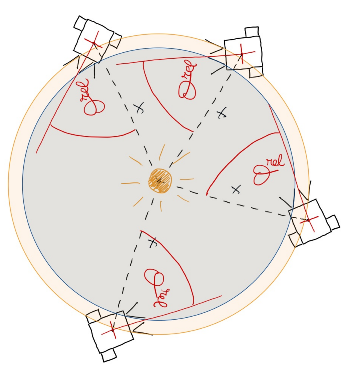

Figure 2: Braitenberg vehicle 2b and what it can distinguish: this vehicle is insensitive to its position on isocontours, or level sets, centred around the stimulus, provided that the signed relative heading θ^{rel} is the same. The body of the vehicle is drawn as a rectangle for graphical convenience, but, for simplicity, implemented as a circle in the simulation.

This follows first from an observational symmetry: isotropy of the light field and fixed sensor geometry imply that a common rotation around the source leaves the two sensor readings unchanged. However, observation equality alone is not sufficient for bisimulation. The transition condition established below additionally uses the rotational equivariance of the vehicle dynamics under every wheel-speed action $a \in A$. States with the same distance from the source and the same relative heading are therefore candidates for identification independently of the particular 2b wiring.

To capture this formally, for a state $s = ((x, y), \theta) \in S$, define

$$q_{2b}(s) := \Big(\|(x,y) - (x_{\mathsf{Stim}}, y_{\mathsf{Stim}})\|, \mathrm{wrap}\Big(\pi - \big(\theta - \arctan 2(y - y_{\mathsf{Stim}}, x - x_{\mathsf{Stim}})\big)\Big)\Big). \quad (4)$$

where $\mathrm{wrap} : \mathbb{R} \to (-\pi, \pi]$ denotes the canonical representative modulo 2π, i.e. it maps any angle to the unique angle in $(-\pi, \pi]$ that differs from it by an integer multiple of 2π. Thus, q_{2b} can be read as passing from Cartesian coordinates to source-centred polar coordinates and then forgetting the absolute polar angle (by projecting it to the one-element set, $\Phi \to \{*\}$). In other words, q_{2b} keeps only how far the vehicle is from the source and how much it would need to turn to face it, while ignoring where it sits around the source in absolute Cartesian terms. More formally, q_{2b} is the orbit map for rotations about the source. A rotation by α acts on a state by rotating the vehicle's position about x_{Stim} and adding α to its absolute heading. The observation map is invariant under this rotation, and applying any fixed action $a \in A$ to two states related by the rotation produces successor states related by the same rotation. Consequently, states in the same rotational orbit have the same image under q_{2b}.

Proposition 9. *Under the rotational invariance of the observation map and the rotational equivariance of the one-step dynamics for every $a \in A$, the kernel relation of q_{2b} is a bisimulation equivalence on the POMDP of Example 7.*

Proof sketch. Let $B_{2b} \subseteq S \times S$ be the relation defined by

$$(s, s') \in B_{2b} \iff q_{2b}(s) = q_{2b}(s'). \quad (5)$$

This is the kernel relation of q_{2b}, i.e. it identifies states that have the same reduced description under q_{2b}.

Now suppose $(s, s') \in B_{2b}$. Then s and s' have the same distance from the source and the same relative heading. Equivalently, they differ only by an overall rotation around the source. By isotropy of the light field and fixed sensor geometry in body coordinates, such a rotation does not change the distances of the sensors from the source, and therefore does not change the sensory readings. Hence the observation map takes the same value on s and s'.

Now fix an action $a \in A$. The one-step updates of the vehicle are compatible with the same rotation: changing only the absolute polar angle rotates the translational increment by the same amount, while the angular increment depends only on the wheel velocities and body geometry, not on the absolute polar angle. Therefore the successors of s and s' again have the same image under q_{2b}, and hence still belong to the same B_{2b}-equivalence class. Thus equivalent states have identical observations and, for every action, transition to the same quotient class. Therefore B_{2b} is a bisimulation equivalence. □

We use this to define the habitat of vehicle 2b.

Example 10 (A Braitenberg vehicle's habitat)**.** Following Definition 5, we define the habitat as the POMDP given by:

- the state space $S' := R \times \Theta^{\mathsf{rel}}$,
- the action space $A := V_l \times V_r$,
- the observation space $O := O_l \times O_r$,
- deterministic transition dynamics given by the product map

$$\begin{aligned} f_R &: R \times \Theta^{\mathsf{rel}} \times V_l \times V_r \to R, \\ f_{\Theta^{\mathsf{rel}}} &: R \times \Theta^{\mathsf{rel}} \times V_l \times V_r \to \Theta^{\mathsf{rel}}, \end{aligned} \tag{6}$$

- a deterministic observation map given by composing

$$\begin{aligned} f^{r_{\mathsf{Body}}}_{\mathsf{Sensor}} &: R \times \Theta^{\mathsf{rel}} \to R_l \times R_r, \\ f_{\mathsf{Stim}} &: R_l \times R_r \to O_l \times O_r. \end{aligned} \tag{7}$$

This quotient depends on the vehicle type's morphology, observation map, embodied dynamics, and wheel-command interface. Since Proposition 9 establishes equivalence under every action, the resulting habitat is not specific to the contralateral 2b wiring. This means that, while Eq. (3) determines the trajectory realised by vehicle 2b within this habitat, the quotient itself remains valid for other reactive controllers sharing the same embodied interface.

Candidate habitats for some exotic vehicles

In this section we sketch three stylised modifications of vehicle 2b. Their projections are candidate coarse-grainings of $(r, \theta^{\mathrm{rel}})$, not proved quotient POMDPs, because their full observation and transition maps are not specified. We use them here as informal outlines of different possible habitats given the same environment, showing how different bodily constraints in the same isotropic light field can make different aspects of the world behaviourally salient.

Braitenberg sunflower

Our first variation is the *Braitenberg sunflower*, a vehicle whose centre is pinned to the ground, so that it can only reorient itself and cannot translate towards or away from the source. We also assume that the two sensory readings are normalised so that only their directional contrast matters, not their overall magnitude given by the distance from the source. This suggests the projection

$$q_{\mathrm{sun}}(r, \theta^{\mathsf{rel}}) = \theta^{\mathsf{rel}}, \tag{8}$$

whose kernel relation identifies states that agree in signed relative heading while ignoring distance.

This indicates a directional candidate habitat, i.e. configurations with the same signed bearing are identified, while distance from the source is ignored. Since the body is pinned and overall light intensity is normalised away, only reorientation relative to the source remains salient, much like a sunflower tracking the sun by reorientation rather than locomotion.

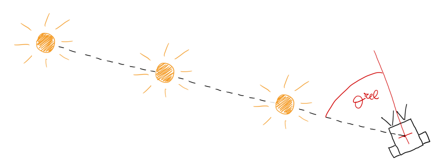

Figure 3: Braitenberg sunflower. This vehicle is pinned at its centre and can only reorient itself. Once overall intensity is normalised away, all lights on the same bearing from the body look the same to it.

Braitenberg sea urchin

Our second variation is the *Braitenberg sea urchin*, having a radially symmetric body with no privileged front, capable of moving directly in any planar direction. Unlike the ordinary vehicle 2b, it does not need to reorient before moving. We assume that its sensing is isotropic, so that what matters is how far the source is from the body, not the angle at which it appears around it. The following projection

$$q_{\mathrm{urchin}}(r, \theta^{\mathsf{rel}}) = r, \tag{9}$$

has kernel relation identifying states that agree in distance from the source while ignoring relative heading.

This suggests a radial candidate habitat in which states at equal source distance are identified. It also requires further assumptions, which we do not develop here. For instance, because of the quantifier over all actions in condition (ii), actions must themselves be rotationally invariant, for example, radial inward and outward actions.

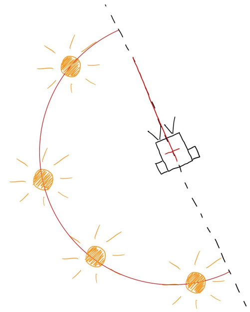

Figure 4: Braitenberg sea urchin. This vehicle is radially symmetric and has no privileged front. With isotropic sensing, all lights at the same distance from the body look the same to it.

Braitenberg crab

Our last variation is the *Braitenberg crab*, a vehicle with a fixed body axis that moves laterally rather than by turning towards the source. For such a body, the central distinction is whether the source lies to the left or to the right of the body. This suggests the coarser projection

$$q_{\mathrm{crab}}(r,\theta^{\mathrm{rel}}) = \begin{cases} L & \text{if } \sin(\theta^{\mathrm{rel}}) > 0, \\ C & \text{if } \sin(\theta^{\mathrm{rel}}) = 0, \\ R & \text{if } \sin(\theta^{\mathrm{rel}}) < 0, \end{cases} \tag{10}$$

whose kernel relation identifies states according to whether the source lies to the left of the body axis, on it, or to the right of it.

This supports the idea of a categorical, lateral candidate habitat that discards distance and angular magnitude while retaining whether the source lies to the left of the body axis, on it, or to its right. This partition is coarser than the sunflower's angular partition, but it is not comparable by refinement with the sea urchin's radial partition. The centre class contains both collinear cases, the source directly ahead and directly behind, which may need to be treated separately in a fuller model. For this projection to induce a quotient POMDP, the transition dynamics must also preserve these three categories under every action.

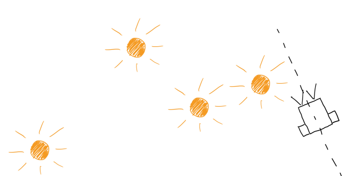

Figure 5: Braitenberg crab. This vehicle moves laterally and responds to the source through a body-centred left–right distinction. All lights on the same side of the body look the same to it.

Taken together, these three examples illustrate a simple point. Even in the same isotropic light field, changing the body's constraints changes the environment that is relevant to the agent. The sunflower foregrounds direction to the source, the sea urchin responds to distance from it, and the crab retains a left-right distinction relative to its body.

Conclusions

In this work, we proposed a formal way of distinguishing the physical environment of a Braitenberg vehicle from its habitat, understood here as the effective environment relevant to its morphology, sensory interface, transition dynamics, and action repertoire. More precisely, we modelled the environment as a POMDP and the habitat as a quotient POMDP induced by a bisimulation equivalence that identifies states with the same observations and equivalent coarse-grained transitions under every action.

For Braitenberg vehicle 2b, this construction shows how a habitat can arise as a coarse-graining of the physical environment. In particular, under the isotropy assumptions adopted here, states that differ only by absolute position around the source while preserving distance and relative heading are identified. This yields a formal example in which a habitat emerges from the agent's lack of sensitivity to some features of the environment while retaining others.

We also introduced three alternative variations of vehicle 2b in the same isotropic light field: a Braitenberg sunflower, a Braitenberg sea urchin, and a Braitenberg crab. These examples sketch some informal candidate habitats illustrating how different bodily constraints can make different features of the world relevant to the agent. The sunflower is sensitive to the direction of the source, but cannot move towards it. The sea urchin is sensitive to the distance from the source, and can move towards it without reorienting. The crab reduces the world further to a body-centred left-right distinction of whether the source lies to the left or to the right of the body axis.

More broadly, the paper offers a formal operationalisation of habitat under a particular agent-environment factorisation. Future work should extend this analysis by giving a fuller formal treatment of these variations and by investigating how a notion closer to umwelt might emerge from habitat so as to capture individual history, learning, and subjective perspective. Furthermore, since the present construction identifies the quotient analytically from a known environmental symmetry, extending the approach proposed here to more complex systems will require integrating existing methods for symmetry detection, model-based partition refinement, and approximate bisimulations.

Acknowledgements

The authors would like to thank Anil Seth, Warrick Roseboom, and Peter Lush for discussing some of these ideas with them. M.B. and F.T. were supported by JST, Moonshot R&D, Grant Number JPMJMS2012.

References

Ay, N. and Löhr, W. (2015). The Umwelt of an embodied agent—a measure-theoretic definition. *Theory in Biosciences*, 134(3-4):105–116.

Baggs, E. and Chemero, A. (2019). The Third Sense of Environment. In *Perception as Information Detection*. Routledge.

Baggs, E. and Chemero, A. (2021). Radical embodiment in two directions. *Synthese*, 198(S9):2175–2190.

Bakirtzis, G., Savvas, M., and Topcu, U. (2025). Categorical Semantics of Compositional Reinforcement Learning. *Journal of Machine Learning Research*, 26.

Baltieri, M., Biehl, M., Capucci, M., and Virgo, N. (2025a). A Bayesian Interpretation of the Internal Model Principle.

Baltieri, M., Torresan, F., and Nakai, T. (2025b). A coalgebraic perspective on predictive processing.

Beer, R. D. (1995). A dynamical systems perspective on agent-environment interaction. *Artificial Intelligence*, 72(1-2):173–215.

Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3):91–99.

Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psychology*. MIT Press.

Brentari, C. (2013). How to make worlds with signs. Some remarks on Jakob von Uexküll's Umwelt theory. *Rivista Italiana di Filosofia del Linguaggio*, 7(2).

Brentari, C. (2015). *Jakob von Uexküll: The Discovery of the Umwelt between Biosemiotics and Theoretical Biology*, volume 9 of *Biosemiotics*. Springer Netherlands, Dordrecht.

Castro, P. S. (2020). Scalable Methods for Computing State Similarity in Deterministic Markov Decision Processes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 10069–10076.

Castro, P. S., Panangaden, P., and Precup, D. (2009). Equivalence Relations in Fully and Partially Observable Markov Decision Processes. In *IJCAI 2009 - Proceedings*.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. MIT Press.

Cronin, T. W. and Bok, M. J. (2016). Photoreception and vision in the ultraviolet. *Journal of Experimental Biology*, 219(18):2790–2801.

Di Paolo, E. A., Buhrmann, T., and Barandiaran, X. (2017). *Sensorimotor Life: An Enactive Proposal*. OUP Oxford, Oxford, first edition.

Feiten, T. E. (2020). Mind After Uexküll: A Foray Into the Worlds of Ecological Psychologists and Enactivists. *Frontiers in Psychology*, 11:480.

Feiten, T. E. (2024). *Jakob von Uexküll's Concept of Umwelt as an Account of the Mental*. PhD thesis, University of Cincinnati.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton, Mifflin and Company.

Givan, R., Dean, T., and Greig, M. (2003). Equivalence notions and model minimization in Markov decision processes. *Artificial Intelligence*, 147(1-2):163–223.

Harvey, I. (1992). Untimed and Misrepresented: Connectionism and the Computer Metaphor. Technical report, University of Sussex, School of Cognitive and Computing Sciences.

Harvey, I. (2008). Misrepresentations. In *Proceedings of the Eleventh International Conference on the Simulation and Synthesis of Living Systems*. MIT Press.

Hepburn, J. and Wonham, W. (1984). Error feedback and internal models on differentiable manifolds. *IEEE Transactions on Automatic Control*, 29(5):397–403.

Hotton, S. and Yoshimi, J. (2024). *The Open Dynamics of Braitenberg Vehicles*. MIT Press.

Jacobs, B. (2017). *Introduction to Coalgebra: Towards Mathematics of States and Observation*. Cambridge University Press, first edition.

Javed, K. and Sutton, R. S. (2024). The big world hypothesis and its ramifications for artificial intelligence. In *Finding the Frame: An RLC Workshop for Examining Conceptual Frameworks*.

Kaelbling, L. P., Littman, M. L., and Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101(1-2):99–134.

Levine, S., Kumar, A., Tucker, G., and Fu, J. (2020). Offline Reinforcement Learning: Tutorial, Review, and Perspectives on Open Problems.

Merleau-Ponty, M. (1942). *La Structure du Comportement*. Presses Universitaires de France, Paris.

Murphy, K. (2025). Reinforcement Learning: An Overview.

Ni, T., Eysenbach, B., Seyedsalehi, E., Ma, M., Gehring, C., Mahajan, A., and Bacon, P.-L. (2024). Bridging State and History Representations: Understanding Self-Predictive RL.

Parr, T., Pezzulo, G., and Friston, K. J. (2022). *Active Inference: The Free Energy Principle in Mind, Brain, and Behavior*. MIT Press.

Philippides, A., Baddeley, B., Cheng, K., and Graham, P. (2011). How might ants use panoramic views for route navigation? *Journal of Experimental Biology*, 214(3):445–451.

Puterman, M. L. (2014). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons.

Rañó, I. (2012). A Model and Formal Analysis of Braitenberg Vehicles 2 and 3. In *2012 IEEE International Conference on Robotics and Automation*.

Ravindran, B. and Barto, A. G. (2003). SMDP Homomorphisms: An Algebraic Approach to Abstraction in Semi-Markov Decision Processes. In *IJCAI'03: Proceedings of the 18th International Joint Conference on Artificial Intelligence*.

Rosas, F., Boyd, A., and Baltieri, M. (2025). AI in a vat: Fundamental limits of efficient world modelling for agent sandboxing and interpretability. In *Proceedings of the Second Reinforcement Learning Conference*, pages 2844–2881.

Rutten, J. J. M. M. (2000a). Coalgebra, Concurrency, and Control. In *Discrete Event Systems*, pages 31–38. Springer US, Boston, MA.

Rutten, J. J. M. M. (2000b). Universal coalgebra: A theory of systems. *Theoretical Computer Science*, 249(1):3–80.

Sangiorgi, D. (2009). On the origins of bisimulation and coinduction. *ACM Transactions on Programming Languages and Systems*, 31(4):1–41.

Shaikh, D. and Rañó, I. (2020). Braitenberg Vehicles as Computational Tools for Research in Neuroscience. *Frontiers in Bioengineering and Biotechnology*, 8.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning Series. The MIT Press, Cambridge, Massachusetts, second edition.

Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Belknap Press.

Van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy*, 92(7):345–381.

Varela, F. J. and Rosch, E. (1992). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press.

von Uexküll, J. (2010). *A Foray into the Worlds of Animals and Humans: With A Theory of Meaning*, volume 12 of *Posthumanities*. University of Minnesota Press, Minneapolis, London.

von Uexküll, J. and Kriszat, G. (1934). *Streifzüge Durch Die Umwelten von Tieren Und Menschen Ein Bilderbuch Unsichtbarer Welten*, volume 21 of *Verständliche Wissenschaft*. Springer, Berlin, Heidelberg.

Wonham, W. M. (1976). Towards an Abstract Internal Model Principle. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(11):735–740.

Zhang, A., Lipton, Z. C., Pineda, L., Azizzadenesheli, K., Anandkumar, A., Itti, L., Pineau, J., and Furlanello, T. (2021a). Learning Causal State Representations of Partially Observable Environments.

Zhang, A., McAllister, R., Calandra, R., Gal, Y., and Levine, S. (2021b). Learning Invariant Representations for Reinforcement Learning without Reconstruction.

Zhang, Y., Luo, Z., and Baltieri, M. (2026). Compositional Behavioral Semantics for State Abstraction in Reinforcement Learning. In *Forty-Third International Conference on Machine Learning*.