CrossMark

The *Umwelt* of an embodied agent—a measure-theoretic definition

Nihat Ay[1,2,3] · Wolfgang Löhr[4]

Abstract We consider a general model of the sensori-motor loop of an agent interacting with the world. This formalises Uexküll's notion of a *function-circle*. Here, we assume a particular causal structure, mechanistically described in terms of Markov kernels. In this generality, we define two σ-algebras of events in the world that describe two respective perspectives: (1) the perspective of an external observer, (2) the intrinsic perspective of the agent. Not all aspects of the world, seen from the external per-spective, are accessible to the agent. This is expressed by the fact that the second σ-algebra is a subalgebra of the first one. We propose the smaller one as formalisation of Uexküll's *Umwelt* concept. We show that, under continuity and compactness assumptions, the global dynamics of the world can be simplified without changing the internal process. This simplification can serve as a minimal world model that the system must have in order to be consistent with the internal process.

✉ Nihat Ay
nay@mis.mpg.de

1 Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, 04103 Leipzig, Germany

2 Faculty of Mathematics and Computer Science, University of Leipzig, PF 100920, 04009 Leipzig, Germany

3 Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

4 Universität Duisburg-Essen, Thea-Leymann-Strasse 9, 45117 Essen, Germany

Introduction: the intrinsic view of embodied agents

Uexküll's *function-circle* and the sensorimotor loop

A key observation based on many case studies within the field of embodied intelligence implies that quite simple control mechanisms can lead to very complex behaviours (Pfeifer and Bongard 2007). This gap between simplicity and complexity related to the same thing, the agent's behaviour, is the result of two different frames of reference in the description. Here, the intrinsic view of the agent, which provides the basis for its control, can greatly differ from the (extrinsic) view of an external observer. This important understanding is not new. In the first half of the last century, Uexküll has conceptualised this understanding by his notion of *Umwelt*, which summarises all aspects of the world that have an effect on the agent and can be affected by the agent (see Von Uexküll 2014). Further-more, he has convincingly exemplified this notion in terms of many biological case studies. These case studies are presented in his book (Von Uexküll 1934), supplemented by insightful illustrations (see in Fig. 1, as an example, the *Umwelt* of a bee).

Uexküll has developed his *Umwelt* concept based on the notion of a *function-circle* (*Funktionskreis*, see Fig. 2a). It graphically represents the causal interaction of an animal with its surroundings. Nowadays this circle is known as the *sensorimotor loop* and it plays an important role within the field of embodied cognition (see Fig. 2b). However, its interpretation has not changed, so that Uexküll's descrip-tion of the *function-circle* perfectly applies to the sensori-motor loop:

"Every animal is a subject, which, in virtue of the structure peculiar to it, selects stimuli from the

2 Springer

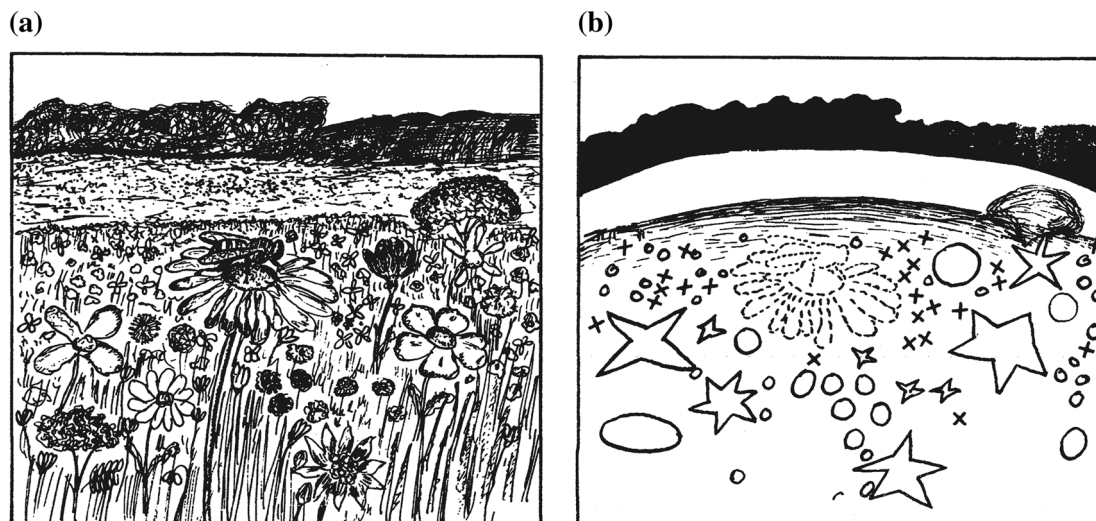**(a)**   **(b)**

Fig. 1 The *Umwelt* of a bee as illustrated in Von Uexküll (1934). **a** The environment of a bee how we perceive it as an external observer. **b** The same bee perceives only particular aspects of the same world, which constitute its *Umwelt*
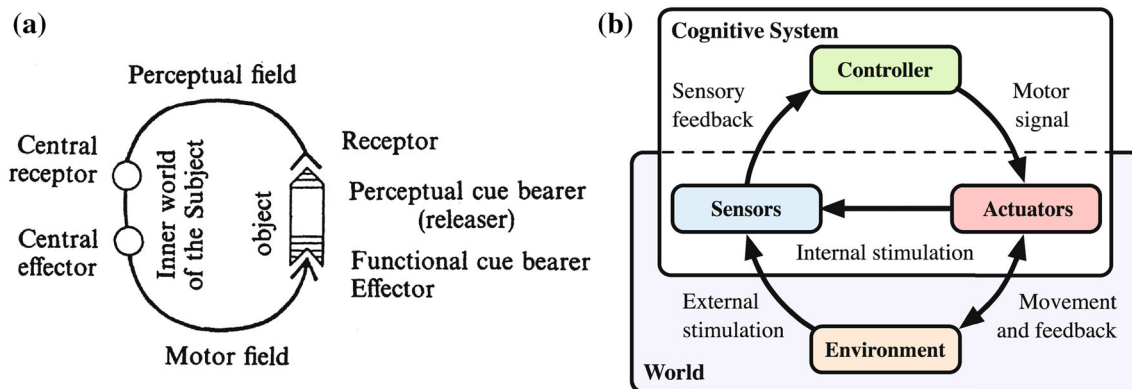
(a)   **(b)**

Fig. 2 **a** Uexküll's *function-circle* (*Funktionskreis*) (Von Uexküll 1934), **b** the sensorimotor loop from the field of embodied cognition (Ay and Ghazi-Zahedi 2014)

general influences of the outer world, and to these it responds in a certain way. These responses, in their turn, consist of certain effects on the outer world, and these again influence the stimuli. In this way there arises a self-contained periodic cycle, which we may call the function-circle of the animal." (Von Uexküll 1926, page 128)

In this section, we introduce the *sensorimotor loop* in terms of a causal diagram (Pearl 2000) of the involved processes, which describes the interaction of the agent with the world (Tishby and Polani 2010; Zahedi et al. 2010). In addition to the causal structure of this interaction, as shown in Fig. 3, we need to formalise the mechanisms that generate the individual processes. A very general way of formalising mechanisms is provided by the notion of a Markov kernel, which is also used in information theory for a mathematical description of a channel. As an example, let

us consider the sensor mechanism denoted by β. Given a state w of the world, the sensor assumes a state s, which can be subject to some noise. Therefore, the sensor output is best described as a probability distribution over the sensor states, that is $\beta(w; ds)$. This reads as the probability that the sensor assumes a state s in the infinitesimal set ds. Equivalently, we can consider the probability $\beta(w; B)$ that the sensor assumes a state s in the set B given the world state w. Clearly, we get this probability of B by integrating the probabilities of the infinitesimal sets ds contained in B: $\beta(w; B) = \int_B \beta(w; ds)$. In addition to the sensor mechanism β we have the actuator mechanism π, referred to as the agent's *policy*. Finally, the mechanisms α and φ describe the dynamics of the world and the agent, respectively. Note that the probabilistic description of the mechanisms does not mean that we exclude deterministic mechanisms. For instance, one might want to assume that the dynamics of

Fig. 3 The causal diagram of the sensorimotor loop. In each instant of time the agent (C) takes a measurement from the world (W) through its sensors (S) and affects the world through its actuators (A)
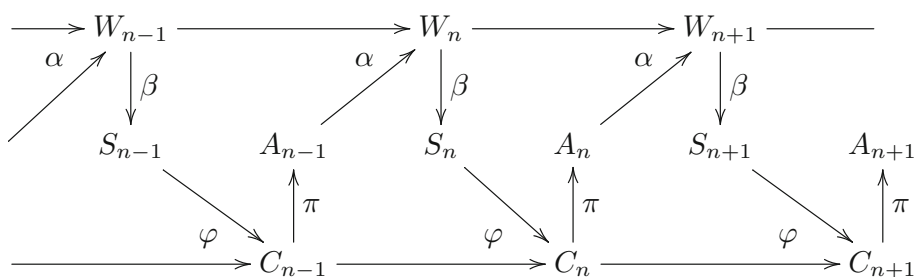
the world is deterministic. Together with an initial distribution of W_0, S_0, C_0, A_0, the Markov kernels α, β, φ, and π, the mechanisms of the sensorimotor loop, specify the distribution of the overall process $W_n, S_n, C_n, A_n, \ n \in \mathbb{N}$, consisting of the individual processes of the world (W), the sensors (S), the agent (C), and the actuators (A).

The causal model of the sensorimotor loop will allow us to formalise what we mean by intrinsic and extrinsic frames of reference in terms of σ-algebras. These are basic mathematical objects from measure theory that naturally describe a set of observables which is assigned to an observer. Already at this very general level, a description of the gap between the intrinsic and extrinsic perspective is possible.

Uexküll's *Umwelt* in terms of a σ-algebra

With this structure at hand, we can formalise Uexküll's agent centric *Umwelt*. Note that the world considered in the sensorimotor loop is meant to be the world as it can be seen from the perspective of an external observer, which is also referred to as *outer world* by Uexküll (Von Uexküll 1926). This perspective is clearly not accessible to the agent. The agent has its own intrinsic view at this world, which implies an agent specific world, consisting only of those objects in the world that the agent can perceive and affect.

Now let us be a bit more precise. Assume that we have two world states w and w' that are distinguishable from the perspective of an external observer. The agent's sensorimotor apparatus, however, may not be rich enough for this distinction. The agent would then perceive w and w' as being the same world state. By this identification, the original set of world states is partitioned into classes that represent the states of the agent's world, its *Umwelt*. This partition is illustrated in Fig. 4.

Let us first assume that we have only finitely many of these classes, say $A_1, A_2, \ldots, A_n$, which represent the internal states of the agent. Given a world state w, the agent will assume one of these classes as internal state, the one that contains w, which we denote by $A(w)$. Now consider an arbitrary subset A of world states. We call A a *distincion* (that the agent can make with its internals states) if the
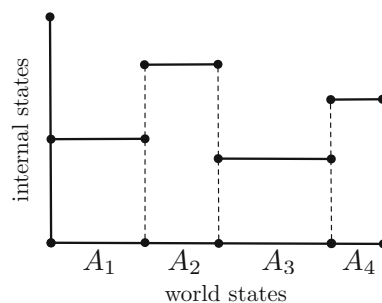

Fig. 4 Clustering of world states

following holds for all world states w: Knowing the internal state $A(w)$ is sufficient to decide whether or not w is contained in A. Clearly, the individual classes themselves are distinctions. But these are not the only ones. For example, $A = A_1 \cup A_2$ is also a distinction. To see this we have to consider three cases: (1) $w \in A_1$: then $w \in A$, (2) $w \in A_2$: then $w \in A$, (3) $w \in A_i$ for some $i > 2$: then $w \notin A$. Thus, if we know to which class w belongs then we know whether or not w is contained in A. More generally, we have the following set $\mathcal{A}$ of distinctions that the agent can make with the internal states $A_1, \ldots, A_n$:

$$\mathcal{A} = \{A_{i_1} \cup A_{i_2} \cup \ldots \cup A_{i_k} \ : \ 1 \le i_1 < i_2 < \ldots < i_k \le n\}. \tag{1}$$

This set is closed under natural operations. Clearly, if A is a distinction, then the complement A^c of A is also a distinction. Furthermore, unions and intersections of distinctions are also distinctions. Having such a set $\mathcal{A}$ of distinctions, we can recover the class to which a world state w belongs by

$$A(w) = [w]_{\mathcal{A}} := \bigcap_{\substack{A \in \mathcal{A} \\ w \in A}} A. \tag{2}$$

This correspondence between partitions into classes and sets of distinctions is one-to-one in the finite case. However, if we drop that assumption the correspondence does not hold anymore, and we have to work with sets of distinctions in the first place, as they encode more information than the corresponding partitions. Extending our reasoning

to possibly infinitely many distinctions, we have to assume that the set $\mathcal{A}$ of distinctions an agent can make with its sensorimotor apparatus satisfies the conditions of a σ-algebra, that is:

1. $\emptyset \in \mathcal{A}$,
2. $A \in \mathcal{A} \;\Rightarrow\; A^c \in \mathcal{A}$,
3. $A_1, A_2, \ldots \in \mathcal{A} \;\Rightarrow\; \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Clearly, as a special case, the set (1) forms a σ-algebra. Given an arbitrary σ-algebra $\mathcal{A}$, we can use formula (2) in order to define the classes that it generates. As already mentioned, this set of classes alone contains less information than the set $\mathcal{A}$ of distinctions. This is the reason why we consider in this article state sets *together* with σ-algebras on these state sets, leading to the notion of so-called measurable spaces.

Note that we do not address the problem of how the sensorimotor apparatus of the agent might instantiate a set of distinctions. In this regard, we want to highlight the following problem of information integration. Let us associate with a distinction A a sensor of the agent that is active if the world state w is in A, and inactive if it is not. Then consider two of such sensors corresponding to the distinctions A_1 and A_2. In principal, knowing that w is in A_1, through the first sensor, and knowing that w is also in A_2, through the second sensor, implies that $w \in A_1 \cap A_2$. But this implication is a purely logical one. It is not clear whether there should be an instance in the system that actually makes the distinction $A_1 \cap A_2$ through a corresponding third sensor. Note also, that we do not assume that the agen's distinctions imply any kind of conscious experiences.

In our motivation of σ-algebras as the right model for describing the intrinsic perspective of an agent we did not explicitly specify the mapping from the world states to internal states. Clearly, this has to be done based on the formal model of the sensorimotor loop as shown in Fig. 3. In order to explain how we are going to use σ-algebras in the context of the sensorimotor loop, consider, for instance, the sensor mechanism β, which generates a sensor state s, given a world state w (Note that β is meant to incorporate all sensors of the agent, not only one.) In general, s will contain information about the world state, which will allow the agent to distinguish it from other world states. One can assign to β a set of distinctions, a σ-algebra, that describes the world as seen through the immediate response s of the sensor β. This σ-algebra is denoted by $\sigma(\beta)$ and referred to as the σ-algebra generated by β. Note, however, that $\sigma(\beta)$ does not contain all distinctions that the agent can make. There are further distinctions, mediated through the time evolution, which incorporates the actuator process of the agent. Therefore, the distinctions that we are going to study are based on both, the sensors and actuators of the agent. The intention

of the article is to define a σ-algebra of distinctions in the world that describe Uexküll's notion of an agent centric world, the agent's *Umwelt*. In what follows, we address the following two natural problems:

1. Which structure in the world is used by the mechanisms of the sensorimotor loop?
2. Which structure of the world is visible from the intrinsic perspective of the agent?

We will show that these problems can be appropriately addressed by defining corresponding σ-algebras. The nature of the main results requires some technical knowledge. We assume basic knowledge from measure and probability theory and refer to the comprehensive volumes (Bogachev 2007a, b) on measure theory and to the textbook (Bauer 1996) on probability theory. However, the technical part will be complemented by an extended summary and conclusions section on the results and how they relate to Uexküll's work. The reader interested in the results at a less formal level might want to first read "Summary and conclusions".

Minimal σ-algebras of the world

Minimal separately measurable σ-algebra

We consider measurable spaces W, S, C, and A as state spaces of the sensorimotor process W_n, S_n, C_n, A_n, $n \in \mathbb{N}$. For technical reasons, we assume that these are Souslin spaces, equipped with their respective Borel σ-algebras $\mathfrak{B}(\mathsf{W})$, $\mathfrak{B}(\mathsf{S})$, $\mathfrak{B}(\mathsf{C})$, and $\mathfrak{B}(\mathsf{A})$. In the Appendix, we collect a few general results that are used in this article, thereby also highlighting the special role of Souslin spaces. In order to address the above problems, we fix the Borel σ-algebras on the "agent part"—S, C, and A—of the system. Based on these internal σ-algebras, we consider various sub-σ-algebras of the Borel σ-algebra on W that describe agent related events in the world. In measure-theoretic terms, we study minimal σ-algebras on W that satisfy natural measurability conditions. The most natural ansatz is given by the distinctions that are possible only through sensor measurements. They correspond to the σ-algebra generated by the kernel β, that is $\sigma(\beta)$. However, this is not necessarily consistent with the world dynamics given by the Markov kernel α. Therefore, we consider the following measurability condition. We call a σ-algebra $\mathcal{W} \subseteq \mathfrak{B}(\mathsf{W})$ **(jointly) measurable** if both β and α remain measurable when W is equipped with $\mathcal{W}$ instead of $\mathfrak{B}(\mathsf{W})$. By general assumption, $\mathfrak{B}(\mathsf{W})$ is jointly measurable. It turns out that joint measurability is a quite strong condition. Therefore, it is natural to consider the following weaker measurability condition. We call $\mathcal{W}$ **separately measurable** if, when we

equip W with $\mathcal{W}$, β is measurable and α is separately measurable in the sense that for every $a \in A$, the Markov kernel

$$\alpha_a : (W, \mathcal{W}) \to \mathcal{P}(W, \mathcal{W}), \qquad w \mapsto \alpha(a, w),$$

is measurable ("$\mathcal{P}$" denotes the set probability distributions, here on the measurable space $(W, \mathcal{W})$). Note that, because α is Borel measurable, the functions $a \mapsto \alpha(a, w)$ are measurable for any $\mathcal{W} \subseteq \mathfrak{B}(W)$.

It is straight-forward to construct the unique minimal (w.r.t. partial ordering by inclusion) separately measurable sub-σ-algebra $\mathcal{W}_{\text{ext}}$ of $\mathfrak{B}(W)$.

Lemma 1 *Let $\mathcal{W}_0 := \sigma(\beta)$ and for $n \in \mathbb{N}$ define $\mathcal{W}_n$ recursively by*

$$\mathcal{W}_n := \sigma(\alpha_a : W \to \mathcal{P}(W, \mathcal{W}_{n-1}), \ a \in A)$$
$$= \sigma(\alpha_a(\cdot\,; B), \ a \in A, B \in \mathcal{W}_{n-1})$$

Then $\mathcal{W}_{\text{ext}} := \sigma\big(\bigcup_{n \in \mathbb{N}} \mathcal{W}_n\big)$ is the unique minimal separately measurable σ-algebra, i.e.

$$\mathcal{W}_{\text{ext}} = \bigcap \big\{ \mathcal{W} \subseteq \mathfrak{B}(W) \,|\, \mathcal{W} \ \sigma\text{-algebra}, \ \beta \ \text{measurable},$$
$$\alpha_a \mathcal{W}\text{-}\mathcal{W}\text{-measurable for all } a \in A \big\}$$

Proof "$\subseteq$": Clearly, any σ-algebra $\mathcal{W}$ from the set on the right-hand side has to contain $\mathcal{W}_0$ because β is $\mathcal{W}$-measurable. Further, if it contains $\mathcal{W}_{n-1}$, it also has to contain $\mathcal{W}_n$, because α_a must be measurable for all $a \in A$. Thus it contains $\mathcal{W}_{\text{ext}}$.

"$\supseteq$": We have to show that $\mathcal{W}_{\text{ext}}$ is separately measurable. β is measurable, because $\mathcal{W}_{\text{ext}} \supseteq \mathcal{W}_0 = \sigma(\beta)$. $\bigcup_n \mathcal{W}_n$ is an intersection stable generator of $\mathcal{W}_{\text{ext}}$, thus for measurability of α_a, it is sufficient that $\alpha_a(\cdot\,; B)$ is $\mathcal{W}_{\text{ext}}$- measurable for $B \in \bigcup_n \mathcal{W}_n$. But by definition of $\mathcal{W}_n$, $\alpha_a(\cdot\,; B)$ is $\mathcal{W}_n$-measurable for $B \in \mathcal{W}_{n-1}$. $\square$

Note that there is no reason why α should be jointly measurable when we equip W with $\mathcal{W}_{\text{ext}}$. When we are working with a separately but not jointly measurable σ-algebra, we are rather working with a family $(\alpha_a)_{a \in A}$ of kernels than with a single kernel α. We do not know if a minimal jointly measurable σ-algebra exists in general. The above construction does not work well for α instead of α_a, because we want a product σ-algebra on $A \times W$ and taking products is not compatible with intersections in the sense that $\mathcal{A} \otimes (\mathcal{W} \cap \mathcal{W}') \subsetneqq (\mathcal{A} \otimes \mathcal{W}) \cap (\mathcal{A} \otimes \mathcal{W}')$ in general. Of course, every jointly measurable σ-algebra is separately measurable and thus has to contain $\mathcal{W}_{\text{ext}}$. Also note that $\mathcal{W}_{\text{ext}}$ need not be countably generated, which might cause technical problems when working with $\mathcal{W}_{\text{ext}}$. Next we show that in the "nice case" where $\mathcal{W}_{\text{ext}}$ is countably generated, α is jointly measurable.

Proposition 2 *If $\mathcal{W}_{\text{ext}}$ is countably generated, then it is jointly measurable, and in particular the unique minimal jointly measurable σ-algebra.*

Proof Let $\mathcal{A} := \mathfrak{B}(A)$. We have to show that $f : A \times W \to [0, 1]$, $(a, w) \mapsto \alpha(a, w; B)$ is $(\mathcal{A} \otimes \mathcal{W}_{\text{ext}})$-measurable for arbitrary choice of $B \in \mathcal{W}_{\text{ext}}$. Because $\mathcal{W}_{\text{ext}}$ is countably generated, $\mathcal{A} \otimes \mathcal{W}_{\text{ext}}$ is a countably generated sub-σ-algebra of the Borel σ-algebra of the Souslin space $A \times W$. It follows from Blackwell's theorem (see [Appendix](#)) and the fact that f is Borel measurable that f is $\mathcal{A} \otimes \mathcal{W}_{\text{ext}}$-measurable if and only if it is constant on the atoms of $\mathcal{A} \otimes \mathcal{W}_{\text{ext}}$. The atoms are obviously of the form $\{a\} \times F$, where $a \in A$ and F is an atom of $\mathcal{W}_{\text{ext}}$. Because α_a is measurable w.r.t. $\mathcal{W}_{\text{ext}}$, $f(a, \cdot)$ is constant on the atom F. Thus, f is constant on $\{a\} \times F$ and therefore jointly measurable. $\square$

A simple sufficient condition for $\mathcal{W}_{\text{ext}}$ to be countably generated is that there are only countably many possible actuator states, i.e. A is countable.

Corollary 3 *Let A be countable. Then $\mathcal{W}_{\text{ext}}$ is countably generated and jointly measurable.*

Proof Because $\mathfrak{B}(S)$ is countably generated, $\mathcal{W}_0$ is countably generated. If $\mathcal{W}_{n-1}$ is countably generated, the same holds for $\sigma(\alpha_a : W \to \mathcal{P}(W, \mathcal{W}_{n-1}))$ for any $a \in A$. Because A is countable, $\mathcal{W}_n$ is generated by a countable union of countably generated σ-algebras, thus it is countably generated, and the same holds for $\mathcal{W}_{\text{ext}}$. $\square$

A countably generated, almost jointly measurable σ-algebra in the memoryless case

In this section, we assume that the agent is memoryless, i.e. C_n is conditionally independent of C_{n-1} given S_n. Then we can concatenate the kernels from W to S, from S to C, and from C to A to obtain a new kernel γ from W to A. We then have the following situation, where the C and S components are marginalised (integrated) out.
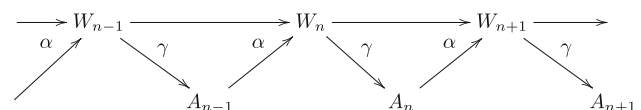
Note that if β is $\mathcal{W}$-measurable, the same holds for γ, but the converse need not be true. We introduce yet another kernel κ which is the combination of γ and α, i.e. $\kappa : W \to \mathcal{P}(A \times W)$, $\kappa(w) = \gamma(w) \otimes \alpha(\cdot, w)$. The reason to do this is that while kernels mapping *from* a product space complicate finding minimal σ-algebras (with product structure), this is not the case for kernels mapping *into* a

product space. The variables $W_n, A_n, n \in \mathbb{N}$, factorise also according to the following graphical model.
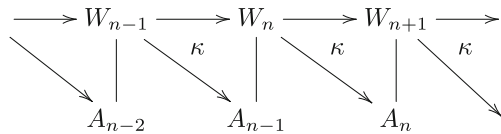
We can define the minimal σ-algebra $\mathcal{W}_\kappa$ such that β and κ are measurable in the same way as we defined $\mathcal{W}_{\mathrm{ext}}$.

Lemma 4 *Let* $\mathcal{W}'_0 := \sigma(\beta)$ *and*

$$\mathcal{W}'_n := \sigma\big(\kappa : W \to \mathcal{P}(A \times W, \mathcal{A} \otimes \mathcal{W}'_{n-1})\big)$$
$$= \sigma\big(\kappa(\,\cdot\,; B),\ B \in \mathcal{A} \otimes \mathcal{W}'_{n-1}\big).$$

Then $\mathcal{W}_\kappa := \sigma\big(\bigcup_{n \in \mathbb{N}} \mathcal{W}'_n\big)$ *is the unique minimal σ-algebra on* W *s.t. β and κ are measurable. Furthermore, $\mathcal{W}_\kappa$ is countably generated.*

Proof Analogous to the proof of Lemma 1 and Corollary 3. □

In the following, consider κ as kernel from $(W, \mathcal{W}_\kappa)$ to $(A \times W, \mathcal{A} \otimes \mathcal{W}_\kappa)$. Note that because $\mathcal{W}_\kappa$ is countably generated, the quotient space obtained by identifying atoms of $\mathcal{W}_\kappa$ to points is again a Souslin space[1] (see Appendix). Thus it is technically nice and, in particular, we can factorise κ into γ and some kernel α' from $A \times W$ to W by choosing regular versions of conditional probability. Then α' is jointly measurable. The draw-back is that $\alpha'(\,\cdot\,, w)$ is only defined $\gamma(w)$-almost surely and we cannot guarantee that $\alpha' = \alpha$ is a valid choice, i.e. that α is $\mathcal{A} \otimes \mathcal{W}_\kappa$-measurable. We easily get the following.

Lemma 5 $\mathcal{W}_\kappa$ *is the unique minimal σ-algebra on* W *that satisfies the following condition. β is measurable and there exists a (jointly measurable) kernel α' from* $A \times W$ *to* W, *s.t., for every measure $\mu \in \mathcal{P}(W, \mathfrak{B}(W))$, $\alpha = \alpha'$ $(\mu \otimes \gamma)$-almost surely.*

Proof The above discussion shows that $\mathcal{W}_\kappa$ satisfies the condition (note that we can w.l.o.g. assume that μ is a Dirac measure). If, on the other hand, α' as above exists, then κ equals the composition of γ and α'. In particular, κ is measurable and Lemma 4 yields the claim. □

The condition $\alpha = \alpha'$ a.s. w.r.t. every measure of the form $\mu \otimes \gamma$ means that the difference between α and α' is not visible regardless of any changes we might impose on the environment. The situation, however, may change if

the agent changes its policy π, thereby changing the kernel γ. Then the difference between α and α' can become important and α' as well as $\mathcal{W}_\kappa$ would have to be changed.

We trivially have that every jointly measurable σ-algebra $\mathcal{W}$ on W must contain $\mathcal{W}_\kappa$. In particular, if $\mathcal{W}_{\mathrm{ext}}$ is countably generated, $\mathcal{W}_\kappa \subseteq \mathcal{W}_{\mathrm{ext}}$. This is probably not true in general.

The world from an intrinsic perspective

Sensory equivalence

In what follows we use equivalence relations to coarse grain the world states and apply the constructions described in Appendix.

Denote by $P_S^{a_\mathbb{N}}(w) \in \mathcal{P}(S^\mathbb{N})$ the distribution of sensor values when the "world" is initially in the state $W_1 = w \in W$ and the agent performs the sequence $a_\mathbb{N} \in A^\mathbb{N}$, of actuator states, i.e. $A_n = a_n$. That is, we modify the agent policy π in a time-dependent way such that π is replaced by $(\pi_n)_{n \in \mathbb{N}}$ and π_n ignores the memory (and thus the sensors) and outputs (the Dirac measure in) the value a_n. The sensor and world update kernels β and α, however, remain unchanged. More explicitly, for $B = B_1 \times \cdots \times B_n \times S \times \cdots \in \mathfrak{B}(S^\mathbb{N})$,

$$P_S^{a_\mathbb{N}}(w_1)(B) = \int \cdots \int \beta(w_1; B_1) \cdots \beta(w_n; B_n)$$
$$\alpha_{a_{n-1}}(w_{n-1}; dw_n) \cdots \alpha_{a_{n-1}}(w_1; dw_2).$$

Now we define an equivalence relation $\sim_s$, called *sensory equivalence*, on W by

$$w \sim_s w' :\Leftrightarrow P_S^{a_\mathbb{N}}(w) = P_S^{a_\mathbb{N}}(w') \quad \forall a_\mathbb{N} \in A^\mathbb{N}.$$

More generally, we obtain the *intrinsic σ-algebra*

$$\mathcal{W}_{\mathrm{int}} := \sigma\big(P_S^{a_\mathbb{N}},\ a_\mathbb{N} \in A^\mathbb{N}\big),$$

which describes the information about the world that can in principle be obtained by the agent through its sensors. Obviously, the atoms $[\cdot]_{\mathcal{W}_{\mathrm{int}}}$ of the intrinsic σ-algebra are given precisely by the sensory equivalence, i.e. $[w]_{\mathcal{W}_{\mathrm{int}}} = \{w' \in W | w' \sim_s w\}$ for all $w \in W$.

It turns out (Proposition 6 and Example 7) that the intrinsic σ-algebra leads to a coarser partitioning of the world than the extrinsic perspective formalised by $\mathcal{W}_{\mathrm{ext}}$. The reason is that the construction of $\mathcal{W}_{\mathrm{ext}}$ uses knowledge of the mechanisms of the "world", more precisely the world update kernel α, which is required to remain measurable when we replace the Borel σ-algebra $\mathfrak{B}(W)$ by $\mathcal{W}_{\mathrm{ext}}$. The necessary information about α cannot be

---

[1] More precisely, there exists a Souslin topology such that the Borel σ-algebra coincides with the final σ-algebra induced by the canonical projection from $(W, \mathcal{W}_\kappa)$ onto the quotient.

constructed from the sensor values in general, even if infinitely many observations are possible and all probabilities can be estimated accurately. The difference between the intrinsic and extrinsic point of view is precisely that the agent does not know the mechanisms of the world encoded in α.

Proposition 6 $\mathcal{W}_{\text{int}} \subseteq \mathcal{W}_{\text{ext}}$. *In particular*

$$[w]_{\mathcal{W}_{\text{ext}}} \subseteq [w]_{\mathcal{W}_{\text{int}}} = \{w' \in \mathsf{W} | w' \sim_s w\} \qquad \forall w \in \mathsf{W}.$$

Furthermore, $\mathcal{W}_{\text{int}} = \mathcal{W}_{\text{ext}}$ if and only if $\mathcal{W}_{\text{int}}$ is separately measurable, i.e. α_a is $\mathcal{W}_{\text{int}}$-$\mathcal{W}_{\text{int}}$-measurable for every $a \in \mathsf{A}$.

Proof β and α_a are measurable w.r.t. $\mathcal{W}_{\text{ext}}$. Because the σ-algebra on $\mathcal{P}(\mathsf{S}^{\mathbb{N}})$ is generated by the evaluations, and the cylinder sets form a generator of the σ-algebra on $\mathsf{S}^{\mathbb{N}}$, measurability of β and α_a implies measurability of the function $P_S^{a_{\mathbb{N}}}$ for every $a_{\mathbb{N}} \in \mathsf{A}^{\mathbb{N}}$. Hence $\mathcal{W}_{\text{int}} \subseteq \mathcal{W}_{\text{ext}}$. This directly implies the corresponding inclusion for the atoms.

If $\mathcal{W}_{\text{int}} = \mathcal{W}_{\text{ext}}$, $\mathcal{W}_{\text{int}}$ is separately measurable by Lemma 1. Conversely, assume that $\mathcal{W}_{\text{int}}$ is separately measurable. Then, again by Lemma 1, $\mathcal{W}_{\text{ext}} \subseteq \mathcal{W}_{\text{int}}$ and hence $\mathcal{W}_{\text{int}} = \mathcal{W}_{\text{ext}}$. $\square$

Equality in the above proposition does not hold in general, as the following example shows.

Example 7   Let $\mathsf{W} := \{1, \ldots, 5\}$, $\mathsf{S} = \{0, 1\}$ and $|\mathsf{A}| = 1$, i.e. the agent is only observing (a state-emitting HMM). Let $\beta(1) = \beta(4) = \beta(5) = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$, $\beta(2) = \delta_0$, and $\beta(3) = \delta_1$. Further let $\alpha(1) = \frac{1}{2}\delta_2 + \frac{1}{2}\delta_3$, $\alpha(2) = \alpha(3) = \alpha(4) = \delta_4$, $\alpha(5) = \delta_5$. α can be illustrated as
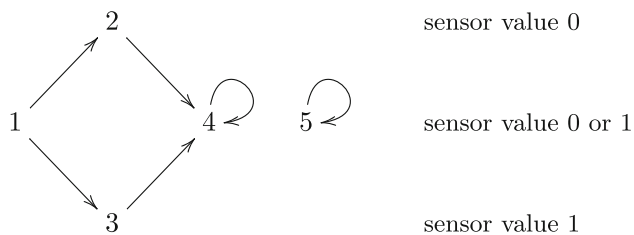
Then $1 \sim_s 4 \sim_s 5 \nsim_s 2, 3$ and $\text{At}(\mathcal{W}_0) = \{\{1, 4, 5\}, \{2\}, \{3\}\}$, but

$$\text{At}(\mathcal{W}_{\text{ext}}) = \text{At}(\mathcal{W}_1) = \{\{1\}, \{2\}, \{3\}, \{4, 5\}\}.$$

Thus 1 and 4 are identified by $\sim_s$ because they produce identical sequences of sensor values, but they are not identified by $\mathcal{W}_{\text{ext}}$ because they have non-identified successors. The definition of $\mathcal{W}_{\text{ext}}$ requires that α remains unchanged, while the same sensor values can be produced with the partition given by $\sim_s$ by modifying α to α', where $\alpha'(1) = \alpha'(4)$ is an arbitrary convex combination of $\alpha(1)$ and $\alpha(4)$.

Sensor-preserving modification of the world

Example 7 suggests that one might be able to interpret the coarser partition given by sensory equivalence as description of the relevant part of the world, provided one is allowed to modify the world update kernel α in such a way that the distribution of sensor values is preserved. Intuitively, one just has to choose one of the values α_a takes on a given $\sim_s$-equivalence class.

Of course these selections have to be done in a measurable way, and we need technical restrictions to deal with this problem. Namely, we assume that the world W is *compact*, and the sensor kernel β as well as the world update kernels $\alpha_a : \mathsf{W} \to \mathcal{P}(\mathsf{W})$ for every given action $a \in \mathsf{A}$ are *continuous*. As usual, $\mathcal{P}(\mathsf{W})$ is equipped with the weak topology induced by bounded continuous functions. Note that compactness and metrisability of W also implies compactness and metrisability of $\mathcal{P}(\mathsf{W})$.

Under these assumptions we can prove that it is possible to modify the world update (and with it the smallest separately measurable σ-algebra $\mathcal{W}_{\text{ext}}$) in such a way that the sensor process is preserved and equality holds in Proposition 6. Furthermore, the "new $\mathcal{W}_{\text{ext}}$" is countably generated and jointly measurable for the modified system.

Definition 8   Let $\alpha' : \mathsf{A} \times \mathsf{W} \to \mathcal{P}(\mathsf{W})$ be an "alternative" world update kernel.

1. We call α' **equivalent** to α if for every $w \in \mathsf{W}$ and $a_{\mathbb{N}} \in \mathsf{A}^{\mathbb{N}}$ the sensor process $P_S^{a_{\mathbb{N}}}(w)$ coincides with the sensor process $P_{S,\alpha'}^{a_{\mathbb{N}}}(w)$ obtained by replacing α with α'.
2. Denote by $\mathcal{W}_{\text{ext}}^{\alpha'}$ the smallest separately measurable σ-algebra of the system where α is replaced by α'.

Remark α' can be seen as a *model* for the mechanisms of the world, which the agent might use. If α' is equivalent to α, α' is a perfect model, as far as the agent's (possible) observations are concerned. Of course it can still make wrong assumptions about aspects of the world that cannot be inferred by the agent. In Proposition 10, we show (under a continuity and compactness assumption) that the agent can always build a perfect model in this sense which is consistent with his intrinsic σ-algebra.

Lemma 9   *Assume that α_a is continuous for every $a \in \mathsf{A}$ and β is continuous. Then $P_S^{a_{\mathbb{N}}}$ is continuous for every $a_{\mathbb{N}} \in \mathsf{A}^{\mathbb{N}}$.*

Proof   Easy to see directly or a special case of (Karr 1975, Thm. 1). $\square$

Proposition 10   *Let W be compact, β and α_a continuous for every $a \in \mathsf{A}$. Then there is a kernel $\alpha' : \mathsf{A} \times \mathsf{W} \to \mathcal{P}(\mathsf{W})$, such that α' is equivalent to α and*

$$\mathcal{W}_{\text{ext}}^{\alpha'} = \mathcal{W}_{\text{int}}. \qquad (3)$$

In particular,

$$[w]_{\mathcal{W}_{\text{ext}}'} = \{w' \in \mathsf{W} \mid w' \sim_s w\} \qquad \forall\, w \in \mathsf{W}.$$

Furthermore, $\mathcal{W}_{\text{ext}}^{\alpha'}$ is countably generated as well as jointly measurable (for the modified system with α replaced by α').

Proof

1. Let $X := \mathcal{P}(\mathsf{S}^{\mathbb{N}})^{\mathsf{A}^{\mathbb{N}}}$ be the set of mappings from action sequences to distributions of sensor sequences, equipped with product topology. Given an initial state w of the world, denote by $F(w)$ the corresponding kernel from action sequences to sensor sequences, i.e. $F : \mathsf{W} \to X$, $F(w) = \big(a_{\mathbb{N}} \mapsto P_S^{a_{\mathbb{N}}}(w)\big)$. Note that F generates $\mathcal{W}_{\text{int}}$, i.e. $\sigma(F) = \mathcal{W}_{\text{int}}$. Because every $P_S^{a_{\mathbb{N}}}$ is continuous (Lemma 9) and X carries the product topology, F is a continuous function from the compact metrisable space W into the Hausdorff space X. In particular, the image $F(\mathsf{W})$ is also compact and metrisable.

2. We can apply a classical selection theorem, e.g. Theorem 6.9.7 in (Bogachev 2007b), and obtain a measurable right-inverse $G : F(\mathsf{W}) \to \mathsf{W}$ with $F \circ G = \text{id}_{F(\mathsf{W})}$. Define $\varsigma := G \circ F$. Then ς is measurable, and $\sigma(\varsigma) \subseteq \sigma(F)$. On the other hand, $\sigma(F) = \sigma(F \circ G \circ F) \subseteq \sigma(\varsigma)$. Hence,

$$\sigma(\varsigma) = \sigma(F) = \mathcal{W}_{\text{int}}. \qquad (4)$$

Define $\alpha'_a := \alpha_a \circ \varsigma$ for every $a \in \mathsf{A}$.

3. A simple induction shows that α' is indeed equivalent to α: For $B = B_1 \times \cdots \times B_n \times \mathsf{S} \times \cdots \in \mathfrak{B}(\mathsf{S}^{\mathbb{N}})$ and $C := B_2 \times \cdots \times B_n \times \mathsf{S} \times \cdots$, we obtain by induction over n

$$P_{S,\alpha'}^{a_{\mathbb{N}}}(w)(B) = \int \beta(w; B_1) P_{S,\alpha'}^{a_{\{2,3,\dots\}}}(\cdot)(C)\, \mathrm{d}\alpha'_{a_1}(w)$$
$$= P_S^{a_{\mathbb{N}}}(\varsigma(w))(B) = P_S^{a_{\mathbb{N}}}(w)(B)$$

4. We claim that $\mathcal{W}_{\text{ext}}^{\alpha'} = \sigma(\varsigma)$. Indeed, α'_a is $\sigma(\varsigma)$-measurable by definition, and β is $\sigma(\varsigma)$-measurable, because $\beta(w)$ is a marginal of $F(w)(a_{\mathbb{N}})$ for any $a_{\mathbb{N}}$. Therefore, $\sigma(\varsigma)$ is separately measurable in the modified system and $\mathcal{W}_{\text{ext}}^{\alpha'} \subseteq \sigma(\varsigma)$. On the other hand, $P_S^{a_{\mathbb{N}}} = P_{S,\alpha'}^{a_{\mathbb{N}}}$ is $\mathcal{W}_{\text{ext}}^{\alpha'}$-measurable, thus the same holds for F and $\sigma(\varsigma) = \sigma(F) \subseteq \mathcal{W}_{\text{ext}}^{\alpha'}$. Hence $\mathcal{W}_{\text{int}} = \mathcal{W}_{\text{ext}}^{\alpha'}$ follows from (4).

5. Since ς is a function into a space with countably generated σ-algebra, $\mathcal{W}_{\text{ext}}^{\alpha'} = \sigma(\varsigma)$ is countably

generated. In particular, it is jointly measurable by Proposition 2.

$\square$

Remark (Non-compact W) For a non-compact world W, we can still obtain an equivalent α' satisfying (3) if we relax the condition that α' needs to be Borel measurable. Instead, it is only universally measurable, i.e. μ-measurable for every $\mu \in \mathcal{P}(\mathsf{W})$. To see this, just replace the selection theorem used in the proof of Proposition 10 by a selection theorem for Souslin spaces, e.g. Theorem 6.9.1 in (Bogachev 2007b). The drawback is that the universal σ-algebra is not countably generated and we do not obtain joint measurability of $\mathcal{W}_{\text{ext}}^{\alpha'}$.

Summary and conclusions

Summary of our definitions and results

We started with the mathematical description of the agent's interaction with the world in terms of a causal diagram. This lead us to the definition of the agent's sensorimotor loop (see Fig. 3), which formalises Uexküll's fundamental notion of a *function-circle*. The sensorimotor loop contains, as part of the description, a reference world, referred to as the *outer world* by Uexküll. It is considered to be objective in the sense that it sets constraints on the distinctions that *any* observer can make in that world. This is formalised in terms of a "large" σ-algebra which contains all reasonable distinctions (the Borel σ-algebra of the world). We defined sub-σ-algebras $\mathcal{W}_{\text{ext}}$ and $\mathcal{W}_{\text{int}}$ that represent two agent specific perspectives. The first one, introduced in "Minimal separately measurable σ-algebra", is based on two requirements:

1. First, we assume that $\mathcal{W}_{\text{ext}}$ contains the distinctions in the world that the agent can make based on the immediate response of its sensors. As the corresponding mechanism is encoded by the Markov kernel β, this means that the σ-algebra generated by β should be contained in $\mathcal{W}_{\text{ext}}$, that is $\sigma(\beta) \subseteq \mathcal{W}_{\text{ext}}$ (see Lemma 1). These distinctions seem to be closely related to Uexküll's *world-as-sensed* (translation of the original term *einfache Merkwelt* (Von Uexküll 1926), page 132). In addition to these aspects of the world, the agent is capable of making also mediated or distal distinctions. We believe that the mediated distinctions correspond to those aspects of the world that Uexküll describes as *the higher grades of the world-as-sensed* (translation of the original *term höhere Stufen der*

Merkwelt (Von Uexküll 1926), page 140). In order to incorporate these mediated distinctions, we impose the next condition on $\mathcal{W}_{\text{ext}}$.

2. With this second condition we basically assume that the system is closed in the sense that all mediated distinctions in the world are taken into account. This is formalised by the iteration formula for $\mathcal{W}_n$ in Lemma 1. The main insight of this lemma is that the closedness is equivalently expressed by the invariance condition

$$\alpha_a^{-1}(\mathcal{W}) \subseteq \mathcal{W}, \qquad a \in \mathsf{A}. \tag{5}$$

Stated differently, incorporating all mediated distinctions, based on an initial set of prime distinctions, is equivalent to enlarging this set until the condition (5) is satisfied. Our σ-algebra $\mathcal{W}_{\text{ext}}$ is then the smallest σ-algebra $\mathcal{W}$ that contains $\sigma(\beta)$ and satisfies this invariance. Condition (5) is related to the closedness of dynamical systems studied in (Pfante and Ay 2015; Pfante et al. 2014).

The definition of $\mathcal{W}_{\text{ext}}$ is quite natural and might appear as the right formalisation of Uexküll's *Umwelt* in terms of a σ-algebra of distinctions. However, the invariance condition requires knowledge about the mechanisms of the world, formalised in terms of the Markov kernel α. Therefore, in "Sensory equivalence" we introduced another σ-algebra, which does not require this knowledge and is defined in an intrinsic manner. It is based on the following sensory equivalence relation: We identify two world states w and w' if they induce the same sensor process, given any sequence $a_1, a_2, \ldots$ of actuator states. We define $\mathcal{W}_{\text{int}}$ to be the σ-algebra associated with this relation. It consists of those distinctions the agent can make in the world that involve both the sensors and actuators. In this sense, $\mathcal{W}_{\text{int}}$ takes into account Uexküll's *perceptual world* (*Merkwelt*, referred to as *world-as-sensed* in (Von Uexküll 1926)) and *effector world* (*Wirkwelt*, referred to as *world of action* in (Von Uexküll 1926)). We do not, however, separate these two worlds as they are intertwined and define, together, $\mathcal{W}_{\text{int}}$.

> "We no longer regard animals as mere machines, but as subjects whose essential activity consists of perceiving and acting. We thus unlock the gates that lead to other realms, for all that a subject perceives becomes his *perceptual world* and all that he does, his *effector world*. Perceptual and effector worlds together form a closed unit, the *Umwelt*." (Von Uexküll 1934, page 320)

In "On decompositions of the Umwelt into Merkwelt and Wirkwelt" below, we comment on other ways to combine the *perceptual world* and the *effector world*. But

first, let us address the following question: How does $\mathcal{W}_{\text{int}}$ relate to $\mathcal{W}_{\text{ext}}$? Their relation is actually quite interesting. First of all, as one would like to have, $\mathcal{W}_{\text{int}} \subseteq \mathcal{W}_{\text{ext}}$, which is the main content of Proposition 6. This means that one can attribute more distinctions to the agent, if the mechanism α of the world is known. Or, stated differently, the agent can operate on the basis of distinctions that are not internally indentifiable in terms of its sensors and actuators. Note that there is one apparent limitation of $\mathcal{W}_{\text{int}}$: It is not invariant in the sense of (5), which means that it does not describe a closed system. On the other hand, making it invariant by extending it sufficiently already leads to our previous σ-algebra $\mathcal{W}_{\text{ext}}$. However, the violation of (5) can only be seen from outside. The agent has no access to the mechanism α from its intrinsic perspective, and it simply does not see whether or not (5) is statisfied. It is quite surprising that, according to our Proposition 10, it is possible for the agent to imagine a mechanism α' that is compatible with $\mathcal{W}_{\text{int}}$ in the sense that it satisfies (5) where α is replaced by α'. With this modification of the mechanism, we have $\mathcal{W}_{\text{int}} = \mathcal{W}_{\text{ext}}^{\alpha'}$, where $\mathcal{W}_{\text{ext}}^{\alpha'}$ is defined in the same way as $\mathcal{W}_{\text{ext}}$ but with the mechanism α' instead of α. Even if the agent is actually operating on the basis of distinctions that are not identifiable from its intrinsic perspective, it is always possible to imagine different mechanisms that only involve the identifiable distinctions. This is why we think that $\mathcal{W}_{\text{int}}$ is the right object for describing the *Umwelt* of an agent.

On decompositions of the *Umwelt* into *Merkwelt* and *Wirkwelt*

Our approach starts with a set of prime distinctions and extends this set by taking into account mediated distinctions that are generated in terms of the actuators of the agent. This way, the *perceptual world* (*Merkwelt*) and the *effector world* (*Wirkwelt*) are incorporated into the *Umwelt* in an asymmetric manner. One could also try to define both worlds separately and integrate them in a symmetric way. Our impression is, however, that this approach has limitations, which we are going to briefly explain.

As argued above, $\mathcal{W}_{\text{merk}} := \sigma(\beta)$ already models the *world-as-sensed*, which we consider, for the moment, to be the same as the *perceptual world*, as both terms are translations of *Merkwelt* (note, however, that *world-as-sensed* denotes its simple form, *einfache Merkwelt*). Constructing a corresponding *effector world* as σ-algebra on W appears less natural. To see this, compare in Fig. 3 the causal link from the world state W_n to the sensor state S_n with the causal link from the actuator state A_{n-1} to the world state W_n. One is directed away from and one toward

W_n. Generally, it is natural to "pull back" distinctions, as we did for the sensor kernel. But without further assumptions there is no natural way to "push forward" distinctions, which would be required for the definition of an *effector world* as σ-algebra on W. Instead of specifying these assumptions and discussing related technical problems, let us simply assume that we already have both, the *perceptual world* $\mathcal{W}_{\text{merk}}$ and the *effector world* $\mathcal{W}_{\text{wirk}}$. How would one combine them so that they "form a closed unit" as Uexküll describes it in the above quote? There are two natural ways to combine the two σ-algebras: The intersection $\mathcal{W}_{\cap} := \mathcal{W}_{\text{merk}} \cap \mathcal{W}_{\text{wirk}}$, which consists of all distinctions, that are contained in both worlds, and the union, $\mathcal{W}_{\vee} := \sigma(\mathcal{W}_{\text{merk}} \cup \mathcal{W}_{\text{wirk}})$, which is the smallest set of distinctions so that both worlds are contained in it. We argue that both choices are limited by applying them to two special cases of the senserimotor loop: In case one, we assume that the agent can only observe but not act, which we refer to as a *passive observer*. Any reasonable definition of the σ-algebra $\mathcal{W}_{\text{wirk}}$ should lead to a trivial *effector world* in this case, that is $\mathcal{W}_{\text{wirk}} = \{\emptyset, \mathsf{W}\}$. In case two, we assume that the agent can act but has no sensors to perceive the consequences of its actions. We refer to this second agent as a *blind actor*. For the blind actor, we obviously have a trivial *perceptual world*, that is $\mathcal{W}_{\text{merk}} = \{\emptyset, \mathsf{W}\}$. The following table summarises the various resulting *Umwelten* for these two cases:

	$\mathcal{W}_{\cap}$	$\mathcal{W}_{\vee}$	$\mathcal{W}_{\text{int}}$
Passive observer	No *Umwelt*	Equals *Merkwelt*	Contains *Merkwelt*
Blind actor	No *Umwelt*	Equals *Wirkwelt*	No *Umwelt*

Let us discuss and compare these simple but instructive outcomes. First, we see that for both, the passive observer and the blind actor, the intersection $\mathcal{W}_{\cap}$ is the trivial σ-algebra $\{\emptyset, \mathsf{W}\}$, which we interpret as a trivial or no *Umwelt*. We argue, however, that one should attribute a non-trivial *Umwelt* to a passive observer, if the observed world is rich enough. The reason is that the agent is capable of making distinctions based on its sensors, without the involvement of its actuators. How about the union $\mathcal{W}_{\vee}$? In this case, we argue that the outcome for the blind actor is not satisfying. A blind actor generates effects in the outer world which are, in principle, visible from the perspective of an external observer in terms of his own set of distinctions. But the blind actor himself has no instantiation of these distinctions made from outside. Would one still attribute these distinctions to the *Umwelt* of the blind actor? We think that this should not be the

case. Finally, as already stated, our definition $\mathcal{W}_{\text{int}}$ treats the two worlds in an asymmetric way. This asymmetry is also expressed by the fact that the passive observer has a possibly non-trivial *Umwelt* (the distinctions are made by the passive observer himself), whereas the blind actor has a trivial one (the effects are visible only in terms of distinctions made by an external observer). Our way of integrating sensing and acting in conceptually different from the way sketched in this section and can be summarised as follows: The prime object is the *perceptual world*! Starting with an initial *perceptual world (world-as-sensed)*, the agent can generate more and more distinctions by utilising its actuators. Thereby, the *perceptual world* is gradually enlarged until it incorporates all distinctions of the agent's *Umwelt* (higher grades of the world-as-sensed).

How to treat the case of multiple agents

We conclude with a rough description of how one can study multiple agents based on the developed tools of this article. For this, we have to couple the individual sensorimotor loops, which Uexküll beautifully describes as follows:

> "The function-circles of the various animals connect up with one another in the most various ways, and together form the function-world of living organisms, within which plants are included. For each individual animal, however, its function-circles constitute a world by themselves, within which it leads its existence in complete isolation." (Uexküll 1926, page 126)

We sketch our ideas on this subject by considering the case of only two agents. The diagram in Fig. 5 shows how the sensorimotor loops of two agents are intertwined. Note that each agent i has its own mechanisms $\beta^{(i)}$, $\varphi^{(i)}$, $\pi^{(i)}$, except that there is only one mechanism α which governs the transitions of the common world given the actuator states of both agents. Let us first take the perspective of agent one. From this perspective, the outer world includes agent two, that is the variable $W_n^{(1)}$ contains W_n, $S_n^{(2)}$, $C_n^{(2)}$, and $A_n^{(2)}$ (see Fig. 6a). The perspective of agent two is symmetric. Here, the outer world $W_n^{(2)}$ of agent two includes the corresponding variables W_n, $S_n^{(1)}$, $C_n^{(1)}$, and $A_n^{(1)}$ (see Fig. 6b). Obviously, in principle the two outer worlds are not contained in each other and, in particular, they are not identical. Furthermore, they share the process W which can be considered as a common world of the two agents.

We can express the fact that the two outer worlds are different more formally by

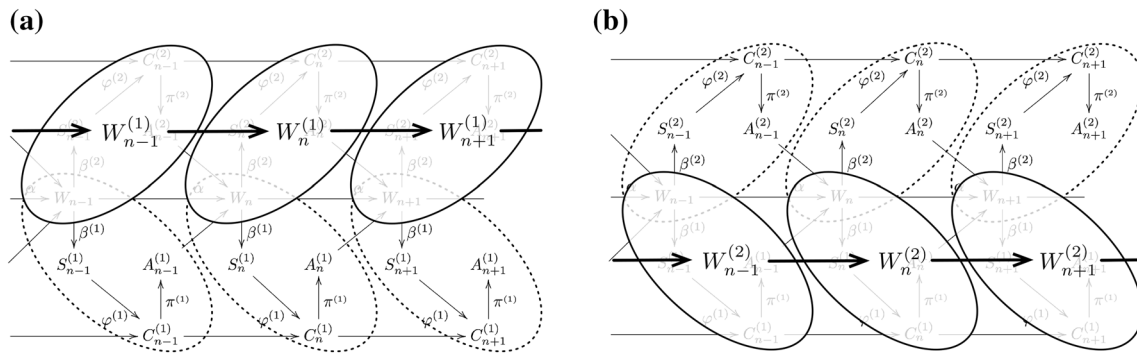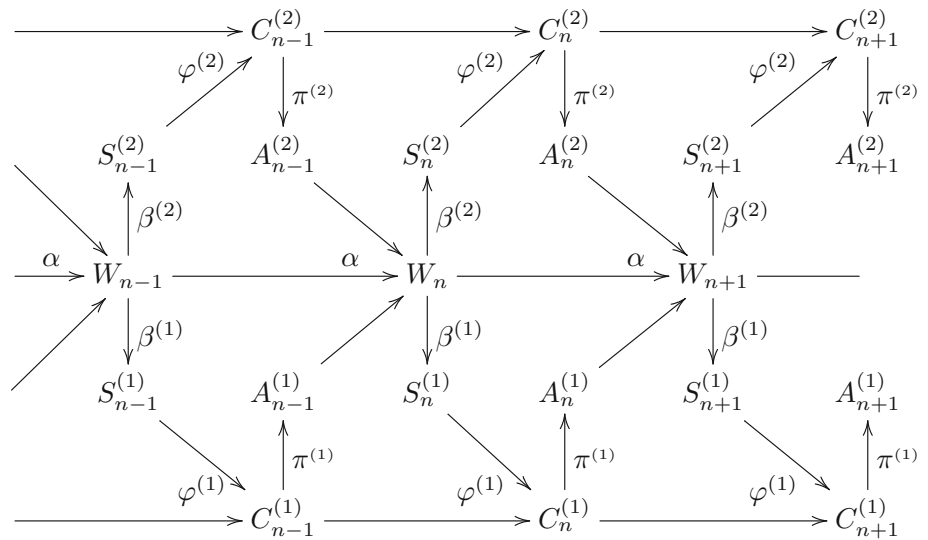Fig. 5 The sensorimotor loops of two agents

Fig. 6 Overlapping but distinct outer worlds of two agents. **a** Outer world of agent one which includes inner world of agent two, and **b** outer world of agent two which includes inner world of agent one

$$\mathfrak{B}(\mathsf{W}^{(1)}) = \mathfrak{B}(\mathsf{W}) \otimes \mathfrak{B}(\mathsf{S}^{(2)}) \otimes \mathfrak{B}(\mathsf{C}^{(2)}) \otimes \mathfrak{B}(\mathsf{A}^{(2)}), \quad (6)$$

$$\mathfrak{B}(\mathsf{W}^{(2)}) = \mathfrak{B}(\mathsf{W}) \otimes \mathfrak{B}(\mathsf{S}^{(1)}) \otimes \mathfrak{B}(\mathsf{C}^{(1)}) \otimes \mathfrak{B}(\mathsf{A}^{(1)}), \quad (7)$$

and we have $\mathcal{W}_{\mathrm{int}}^{(i)} \subseteq \mathcal{W}_{\mathrm{ext}}^{(i)} \subseteq \mathfrak{B}(\mathsf{W}^{(i)})$, $i = 1, 2$. The intersections of these sets of distinctions will be contained in the common world, that is

$$\mathcal{W}_{\mathrm{int}}^{(1)} \cap \mathcal{W}_{\mathrm{int}}^{(2)} \ \subseteq \ \mathcal{W}_{\mathrm{ext}}^{(1)} \cap \mathcal{W}_{\mathrm{ext}}^{(2)} \ \subseteq \ \mathfrak{B}(\mathsf{W}). \quad (8)$$

This highlights an important point. The two agents can share some distinctions in the world. However, these have to be contained in the common world. Each agent can make further distinctions in its respective outer world, so that the *Umwelten* will be in general different. If we consider a probability measure, however, then we can actually identify distinctions $A_1 \in \mathcal{W}_{\mathrm{int}}^{(1)}$ of agent one with distinctions $A_2 \in \mathcal{W}_{\mathrm{int}}^{(2)}$ of agent two if their inter-

section $A_1 \cap A_2$ has full probability. This way, the two agents can, in principle, synchronise and reach some consensus on their respective intrinsic worlds. The more generic situation will be, that the intrinsic worlds are similar rather than perfectly identical with respect to the underlying probability measure. In order to quantify how close the intrinsic worlds $\mathcal{W}_{\mathrm{int}}^{(i)}$, the *Umwelten*, of individual agents are, appropriate distance measures for σ-algebras will be required. Such measures have been studied in the probability theory and statistics literature (Boylan 1971; Neveu 1972; Rogge 1974), and might be applicable to the present context. However, it is not within the scope of this article to present and discuss these measures.

Appendix: state reduction and quotient construction

Let $(X, \mathcal{F})$ be a measurable space. The **atom** of $\mathcal{F}$ containing $x \in X$ and the set of atoms of $\mathcal{F}$ are defined as

$$[x]_{\mathcal{F}} := \bigcap_{x \in F \in \mathcal{F}} F \quad \text{and} \quad \text{At}(\mathcal{F}) := \big\{ [x]_{\mathcal{F}} | x \in X \big\}.$$

Note that if $\mathcal{F}$ is countably generated, $[x]_{\mathcal{F}}$ is a measurable set, $[x]_{\mathcal{F}} \in \mathcal{F}$. In general, however, $[x]_{\mathcal{F}}$ need not be measurable. We recall Blackwell's theorem.

Blackwell's theorem *Let X be a Souslin space and $\mathcal{F} \subseteq \mathfrak{B}(X)$ a countably generated sub-σ-algebra of the Borel σ-algebra. Then*

$$\mathcal{F} = \Big\{ F \in \mathfrak{B}(X) | F = \bigcup_{x \in F} [x]_{\mathcal{F}} \Big\}.$$

Corollary 11 *Let X be a Souslin space, $\mathcal{F} \subseteq \mathfrak{B}(X)$ a countably generated σ-algebra, and $f : X \to \mathbb{R}$ measurable. Then f is $\mathcal{F}$-measurable if and only if it is constant on the atoms of $\mathcal{F}$.*

According to Blackwell's theorem, a countably generated sub-σ-algebra of a Souslin space X is uniquely determined by the set of it atoms. $\text{At}(\mathcal{F})$ is a partition of X into $\mathfrak{B}(X)$-measurable sets. Note, however, that not every partition of X into measurable sets is the set of atoms of a countably generated sub-σ-algebra of $\mathfrak{B}(X)$.

Given any measurable space $(X, \mathcal{F})$, we can define the quotient space $X_{\mathcal{F}}$ as the set $\text{At}(\mathcal{F})$ of atoms of $\mathcal{F}$ equipped with the final σ-algebra $\mathcal{X}_{\mathcal{F}}$ of the canonical projection $[\cdot]_{\mathcal{F}} : X \to \text{At}(\mathcal{F})$. Then a set $B \subseteq X_{\mathcal{F}}$ of atoms is by definition measurable iff $\bigcup B = \bigcup_{[x]_{\mathcal{F}} \in B} [x]_{\mathcal{F}} \in \mathcal{F}$. Note that, obviously, $B \mapsto \bigcup B$ is a complete isomorphism of boolean algebras from $\mathcal{X}_{\mathcal{F}}$ onto $\mathcal{F}$. The following lemma follows easily from the standard theory of analytic measurable spaces and is one of the reasons why Souslin spaces, rather than Polish spaces, are the "right" class of spaces to work with in our setting.

Lemma 12 *Let X be a Souslin space and $\mathcal{F} \subseteq \mathfrak{B}(X)$ a sub-σ-algebra. Then $\mathcal{X}_{\mathcal{F}}$ is the Borel σ-algebra of some Souslin topology on $X_{\mathcal{F}}$ if and only if $\mathcal{F}$ is countably generated.*

References

Ay N, Ghazi-Zahedi K (2014) On the causal structure of the sensorimotor loop. In: Prokopenko M (ed) Guided self-organization: inception. Springer, New York

Bauer H (1996) Probability theory. Walter de Gruyter

Bogachev VI (2007) Measure theory, vol I. Springer, New York

Bogachev VI (2007) Measure theory, vol II. Springer, New York

Boylan ES (1971) Equiconvergence of martingales. Ann Math Stat 42:552–559

Karr AF (1975) Weak convergence of a sequence of Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete 33:41–48

Neveu J (1972) Note on the tightness of the metric on the set of complete sub σ-algebras of a probability space. Ann Math Stat 43:1369–1371

Pfante O, Ay N (2015) Operator-theoretic identification of closed subsystems of dynamical systems. Discontinuity Nonlinearity Complex 1:91–109

Pfeifer R, Bongard J (2007) How the body shapes the way we think. MIT Press, Cambridge

Pfante O, Bertschinger N, Olbrich E, Ay N, Jost J (2014) Comparison between different methods of level identification. Adv Complex Syst 17:1450007

Pearl J (2000) Causality: models. Reasoning and inference. Cambridge University Press, Cambridge

Rogge L (1974) Uniform inequalities for conditional expectations. Ann Prob 2:486–489

Tishby N, Polani D (2010) Information theory of decisions and actions. In: Cutsuridis V, Hussain A, Taylor J (eds) Perception-action cycle: models, architecture and hardware, pp 601–636

Von Uexküll J (1926) Theoretical biology. Harcourt, Brace & Co., New York

Von Uexküll J (1934) A stroll through the worlds of animals and men. In: Lashley K (ed) Instructive behavior. International University Press

Von Uexküll J (2014) Umwelt und Innenwelt der Tiere. In: Mildenberger F, Herrmann B (eds) Klassische Texte der Wissenschaft. Springer, New York

Zahedi K, Ay N, Der R (2010) Higher coordination with less control: a result of information maximization in the sensori-motor loop. Adapt Behav 18:338–355