Polynomial Life: the Structure of Adaptive Systems

Toby St. Clere Smithe
Topos Institute
toby@topos.institute

We extend our earlier work on the compositional structure of cybernetic systems in order to account for the embodiment of such systems. All their interactions proceed through their bodies' boundaries: sensations impinge on their surfaces, and actions correspond to changes in their configurations. We formalize this morphological perspective using polynomial functors. The 'internal universes' of systems are shown to constitute an indexed category of statistical games over polynomials; their dynamics form an indexed category of behaviours. We characterize *active inference doctrines* as indexed functors between such categories, resolving a number of open problems in our earlier work, and pointing to a formalization of the *free energy principle* as adjoint to such doctrines. We illustrate our framework through fundamental examples from biology, including homeostasis, morphogenesis, and autopoiesis, and suggest a formal connection between spatial navigation and the process of proof.

1 Introduction

In a submission to ACT 2020 [10], we presented some first steps towards a theory of categorical cybernetics, motivated by concerns about what gives physical systems life. We explained that perception and action could both be described as processes of Bayesian inference: on the one hand, adjusting beliefs about the world on the basis of observational evidence; on the other, adjusting the world itself in order better to match beliefs. In each case, the system must instantiate a number of structures: a choice of 'prior' belief about the state of the world; a mechanism to generate predictions about sense-data on the basis of that belief, called a 'stochastic channel'; and a (typically approximate) Bayesian inversion of that channel, by which to update those beliefs, in light of sensory observations.

The pairing of a prior with a stochastic channel corresponds to what is called in the informal literature a *generative model*, and it is common to suppose that these models are hierarchical: that is, that the stochastic channel factors as some composite; one imagines that each factor corresponds to predictions at some level of detail, cascading for example down from high-level abstractions to individual photoreceptors. In our earlier submission, we formalized this compositional structure using the bidirectional 'lens' pattern[1] — since predictions and inversions are oppositely directed — and characterized a number of approximate inference processes using a novel category of *statistical games*, whose best responses correspond to optimal inferences. A cybernetic system was then defined as a 'dynamical realisation' of such a game.

This formalism left some things to be desired: our notion of dynamical realisation was ill-defined, and the notion of 'action' was overly abstract. In this submission, we resolve these issues, substantially simplifying our presentation along the way. We explain that various recipes for performing approximate inference, corresponding to our earlier informal notion of dynamical realisation, form functorial *approximate inference doctrines*, between appropriate categories of statistical games and dynamical systems. Then, to formalize a satisfactory notion of action, we note that any active system has a boundary defining

[1] See [9] for a pedagogical presentation of the fundamental structures.

Submitted to:
ACT 2021

its morphology, and that it acts by changing the shape of this boundary; in order to *act on* another system, it couples part of this boundary to that other system, so as to change the composite shape.

To formalize the shapes of systems and their interactions, we adopt polynomial functors: each polynomial will encode the 'phenotype' (possible shapes or configurations) of a system, and the sensorium possible in each configuration. To give such systems life, we construct categories of statistical games and dynamical behaviours indexed by polynomials. An *active inference doctrine* is then an indexed functor between such categories. This framework enables a number of possibilities: we can construct a generative model for a corporation on the basis of models for its employees; we give compositional descriptions of fundamental processes of life such as homeostasis and morphogenesis, and point towards an account of autopoiesis; and we sketch the process by which living systems internalize the structures of their environments, and navigate accordingly, noting that such navigation in abstract spaces corresponds precisely to the process of proof.

The work presented here is work in progress, and owing to constraints of space and time, it has not been possible to elaborate everything that we would have liked. We give proofs of the principal novel results in an appendix, and sketch the rest. Notwithstanding these limitations, we believe our results go some of the way to answering the open questions, of elegance and interaction, sketched at the end of our earlier submission. We see this work as making steps towards a theory of *embodied* cybernetics.

Acknowledgements We thank the members of the Topos Institute for stimulating discussions, and the Foundational Questions Institute and Topos Institute for financial support.

2 Simpler Statistical Games

We begin by presenting a refinement of the statistical games formalism developed in [10]. First, we recall the bidirectional structure of Bayesian inversion.

2.1 Bayesian Lenses

Definition 2.1 ([11])**.** We define the category $\mathbf{GrLens}_F$ of **Grothendieck lenses** for a (pseudo)functor $F : \mathscr{C}^{\text{op}} \to \mathbf{Cat}$ to be the total category of the Grothendieck construction for the pointwise opposite of F. Explicitly, its objects $(\mathbf{GrLens}_F)_0$ are pairs (C,X) of objects C in $\mathscr{C}$ and X in $F(C)$, and its hom-sets $\mathbf{GrLens}_F\big((C,X),(C',X')\big)$ are given by dependent sums

$$\mathbf{GrLens}_F\big((C,X),(C',X')\big) = \sum_{f:\mathscr{C}(C,C')} F(C)\big(F(f)(X'),X\big) \tag{1}$$

so that a morphism $(C,X) \nrightarrow (C',X')$ is a pair $(f,f^\dagger)$ of $f : \mathscr{C}(C,C')$ and $f^\dagger : F(C)\big(F(f)(X'),X\big)$. We call such pairs **Grothendieck lenses** for F or F*-lenses*.

Proposition 2.2. The identity Grothendieck lens on (C,X) is $\mathsf{id}_{(C,X)} = (\mathsf{id}_C,\mathsf{id}_X)$. Sequential composition is as follows. Given $(f,f^\dagger) : (C,X) \nrightarrow (C',X')$ and $(g,g^\dagger) : (C',X') \nrightarrow (D,Y)$, their composite $(g,g^\dagger) \mathbin{⦶} (f,f^\dagger)$ is defined to be the lens $\big(g \bullet f, F(f)(g^\dagger)\big) : (C,X) \nrightarrow (D,Y)$. Associativity and unitality of composition follow from functoriality of F. □

Definition 2.3. Let $(\mathscr{C}, \otimes, I)$ be a monoidal category enriched in a Cartesian closed category $\mathbf{V}$. Define the $\mathscr{C}$**-state-indexed** category $\mathsf{Stat} : \mathscr{C}^{\,\mathrm{op}} \to \mathbf{V}\text{-}\mathbf{Cat}$ as follows.

$$
\begin{aligned}
\mathsf{Stat} \;:\; & \mathscr{C}^{\,\mathrm{op}} \to \mathbf{V}\text{-}\mathbf{Cat} \\
X \mapsto \mathsf{Stat}(X) := & \begin{pmatrix} & \mathsf{Stat}(X)_0 & := & \mathscr{C}_0 \\ & \mathsf{Stat}(X)(A,B) & := & \mathbf{V}(\mathscr{C}(I,X), \mathscr{C}(A,B)) \\ \mathsf{id}_A : & \mathsf{Stat}(x)(A,A) & := & \begin{cases} \mathsf{id}_A : \mathscr{C}(I,X) \to \mathscr{C}(A,A) \\ \quad \rho \quad \mapsto \quad \mathsf{id}_A \end{cases} \end{pmatrix}
\end{aligned}
\tag{2}
$$

$$
f : \mathscr{C}(Y,X) \mapsto \begin{pmatrix} \mathsf{Stat}(f) : & \mathsf{Stat}(X) & \to & \mathsf{Stat}(Y) \\ & \mathsf{Stat}(X)_0 & = & \mathsf{Stat}(Y)_0 \\ & \mathbf{V}(\mathscr{C}(I,X), \mathscr{C}(A,B)) & \to & \mathbf{V}(\mathscr{C}(I,Y), \mathscr{C}(A,B)) \\ & \alpha & \mapsto & f^*\alpha : \big(\sigma : \mathscr{C}(I,Y)\big) \mapsto \big(\alpha(f \bullet \sigma) : \mathscr{C}(A,B)\big) \end{pmatrix}
$$

Composition in each fibre $\mathsf{Stat}(X)$ is given by composition in $\mathscr{C}$. Explicitly, given $\alpha : \mathbf{V}(\mathscr{C}(I,X), \mathscr{C}(A,B))$ and $\beta : \mathbf{V}(\mathscr{C}(I,X), \mathscr{C}(B,C))$, their composite is $\beta \circ \alpha : \mathbf{V}(\mathscr{C}(I,X), \mathscr{C}(A,C)) := \rho \mapsto \beta(\rho) \bullet \alpha(\rho)$, where here we indicate composition in $\mathscr{C}$ by $\bullet$ and composition in the fibres $\mathsf{Stat}(X)$ by $\circ$.

Definition 2.4. Instantiating the category of Grothendieck F-lenses $\mathbf{GrLens}_F$ with $F = \mathsf{Stat} : \mathscr{C}^{\,\mathrm{op}} \to \mathbf{V}\text{-}\mathbf{Cat}$, we obtain the category $\mathbf{GrLens}_{\mathsf{Stat}}$ whose objects are pairs (X,A) of objects of $\mathscr{C}$ and whose morphisms $(X,A) \nrightarrow (Y,B)$ are **Bayesian lenses**: elements of the hom objects

$$
\mathbf{GrLens}_{\mathsf{Stat}}\big((X,A),(Y,B)\big) \cong \mathscr{C}(X,Y) \times \mathbf{V}\big(\mathscr{C}(I,X), \mathscr{C}(B,A)\big)\,. \tag{3}
$$

The identity Stat-lens on (Y,A) is $(\mathsf{id}_Y, \mathsf{id}_A)$, where by abuse of notation $\mathsf{id}_A : \mathscr{C}(I,Y) \to \mathscr{C}(A,A)$ is the constant map id_A defined in (2) that takes any state on Y to the identity on A. The sequential composite of $(c, c^\dagger) : (X,A) \nrightarrow (Y,B)$ and $(d, d^\dagger) : (Y,B) \nrightarrow (Z,C)$ is the Stat-lens $\big((d \bullet c), (c^\dagger \circ c^* d^\dagger)\big) : (X,A) \nrightarrow (Z,C)$ with $(d \bullet c) : \mathscr{C}(X,Z)$ and where $(c^\dagger \circ c^* d^\dagger) : \mathbf{V}\big(\mathscr{C}(I,X), \mathscr{C}(C,A)\big)$ takes a state $\pi : \mathscr{C}(I,X)$ on X to the channel $c^\dagger_\pi \bullet \mathrm{d}^\dagger_{c \bullet \pi}$.

Theorem 2.5 ([9])**.** Let $(c, c^\dagger)$ and $(d, d^\dagger)$ be sequentially composable exact Bayesian lenses. Then the contravariant component of the composite lens $(d, d^\dagger) \,\S\, (c, c^\dagger) = (d \bullet c, c^\dagger \circ c^* d^\dagger)$ is, up to $d \bullet c \bullet \pi$-almost-equality, the Bayesian inversion of $d \bullet c$ with respect to any state π on the domain of c such that $c \bullet \pi$ has non-empty support.

2.2 Statistical Games

The performance of a statistical or cybernetic system depends upon its interaction with its environment, and the prior beliefs that it started with. We will therefore define a statistical game to be a Bayesian lens paired with a fitness function measuring performance in context; and for this we need a notion of context.

Definition 2.6. We define the type of **contexts** for a Bayesian lens of type $(X,A) \nrightarrow (Y,B)$ to be

$$
\overline{\mathbf{BayesLens}}_{\mathscr{C}}\big((X,A),(Y,B)\big) := \mathbf{BayesLens}_{\mathscr{C}}\big((I,I),(X,A)\big) \times \mathbf{BayesLens}_{\mathscr{C}}\big((Y,B),(I,I)\big)
$$

where I is the monoidal unit in $\mathscr{C}$.

Proposition 2.7. In the typical case where I is terminal in $\mathscr{C}$, we have

$$\overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Y,B)\big) \cong \mathscr{C}(I,X)\times\mathscr{C}(Y,A)\,.$$

The proof is a straightforward calculation and we omit it.

Proposition 2.8. Given a context $(\pi,k) : \overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Z,C)\big)$ for a composite lens $(X,A) \overset{f}{\nrightarrow} (Y,B) \overset{g}{\nrightarrow} (Z,C)$, we can obtain contexts for the factors as follows. We have

$$(\pi, k \circ g) : \overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Y,B)\big)\,,$$
$$(f \circ \pi, k) : \overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((Y,B),(Z,C)\big)\,.$$

We are now in a position to define the category of statistical games over $\mathscr{C}$.

Proposition 2.9. Let $\mathscr{C}$ be a category admitting Bayesian inversion. There is a category $\mathbf{SGame}_{\mathscr{C}}$ whose objects are the objects of $\mathbf{BayesLens}_{\mathscr{C}}$ and whose morphisms $(X,A)\to(Y,B)$ are **statistical games**: pairs (f,ϕ) of a lens $f : \mathbf{BayesLens}_{\mathscr{C}}\big((X,A),(Y,B)\big)$ and a fitness function $\phi : \mathbf{Set}\big(\overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Y,B)\big),\mathbb{R}\big)$.

Proof. Deferred to §A.1. □

Definition 2.10. We will write $\mathbf{SimpSGame}_{\mathscr{C}} \hookrightarrow \mathbf{SGame}_{\mathscr{C}}$ for the full subcategory of $\mathbf{SGame}_{\mathscr{C}}$ defined on simple Bayesian lenses $(X,X) \nrightarrow (Y,Y)$. Since duplicating the objects in the pairs (X,X) is redundant, we will write the objects simply as X and the morphisms as $X \nrightarrow Y$ accordingly.

We now present some key examples of statistical games.

Example 2.11. A **simple Bayesian inference game** is any game whose lens is simple $(c,c') : Z \nrightarrow X$ and whose loss function is $\mathbb{E}_{x\sim k\bullet c\bullet\pi}\Big[D_{KL}(c'_\pi(x), c^\dagger_\pi(x))\Big]$, where $\mathbb{E}$ denotes expectation, $\pi : I \to Z$, $k : X \to X$, D_{KL} denotes the Kullback-Leibler divergence, and $c^\dagger$ is the exact inversion of c.

Typically, computing $D_{KL}(c'_\pi(x), c^\dagger_\pi(x))$ is computationally difficult, so one resorts to optimizing an upper bound; a prominent choice is the *free energy*.

Definition 2.12. Let (π,c) be a generative model with $c : Z \to X$. Let $p_c : X\times Z \to \mathbb{R}_+$ and $p_\pi : Z \to \mathbb{R}_+$ be density functions corresponding to c and π. Let $p_{c\bullet\pi} : X \to \mathbb{R}_+$ be a density function for the composite $c\bullet\pi$. Let c'_π be a channel $X \to Z$ that we take to be an approximation of the Bayesian inversion of c and that admits a density function $q : Z\times X \to \mathbb{R}_+$. The **free energy** of c'_π with respect to the generative model given an observation $x : X$ is the quantity

$$\phi(x) := \mathop{\mathbb{E}}_{z\sim c'_\pi(x)}\left[\log\frac{q(z|x)}{p_c(x|z)\cdot p_\pi(z)}\right]. \tag{4}$$

Proposition 2.13. Let D_{KL} denote the Kullback-Leibler divergence between two distributions. The free energy satisfies the following equality:

$$\mathop{\mathbb{E}}_{z\sim c'_\pi(x)}\left[\log\frac{q(z|x)}{p_c(x|z)\cdot p_\pi(z)}\right] = D_{KL}\left[c'_\pi(x), c^\dagger_\pi(x)\right] - \log p_{c\bullet\pi}(x)$$

Since $\log p_{c\bullet\pi}(x)$ is always negative, the free energy is an upper bound on $D_{KL}\Big[c'_\pi(x), c^\dagger_\pi(x)\Big]$.

Proof. Let $p_\omega : Z \times X \to \mathbb{R}_+$ be the density function $p_\omega(z,x) := p_c(x|z) \cdot p_\pi(z)$ corresponding to the joint distribution of the generative model. We have the following equalities:

$$\begin{aligned}
-\log p_{c\bullet\pi}(x) &= \mathop{\mathbb{E}}_{z\sim c'_\pi(x)} \left[-\log p_{c\bullet\pi}(x)\right] \\
&= \mathop{\mathbb{E}}_{z\sim c'_\pi(x)} \left[-\log \frac{p_\omega(z,x)}{p_{c^\dagger_\pi}(z|x)}\right] \qquad \text{(by Bayes' rule)} \\
&= \mathop{\mathbb{E}}_{z\sim c'_\pi(x)} \left[-\log \frac{p_\omega(z,x)}{q(z|x)} \frac{q(z|x)}{p_{c^\dagger_\pi}(z|x)}\right] \\
&= -\mathop{\mathbb{E}}_{z\sim c'_\pi(x)} \left[\log \frac{p_\omega(z,x)}{q(z|x)}\right] - D_{KL}\left[c'_\pi(x), c^\dagger_\pi(x)\right]
\end{aligned}$$

□

Any system that performs (approximate) Bayesian inversion can thus be seen as minimizing some free energy. The *free energy principle* says that all it means to be an adaptive system is to embody a process of approximate inference in this way. We therefore define free energy games:

Example 2.14. A corresponding **simple free energy game** is any game whose lens is simple (c, c') : $Z \mathrel{+\!\!\!\to} X$ and whose loss function is given by

$$\begin{aligned}
&\mathop{\mathbb{E}}_{x\sim k\bullet c\bullet\pi} \left[\mathop{\mathbb{E}}_{z\sim c'_\pi(x)} \left[-\log p_c(x|z)\right] + D_{KL}(c'_\pi(x), \pi)\right] \\
&= \mathop{\mathbb{E}}_{z\sim c'_\pi\bullet k\bullet c\bullet\pi} \left[-\int_X \log p_c(\mathrm{d}k \bullet c \bullet \pi|z)\right] + D_{KL}(c'_\pi \bullet k \bullet c \bullet \pi, \pi)
\end{aligned}$$

where $\pi : I \mathrel{\bullet\!\!\!\to} Z$ and $k : X \mathrel{\bullet\!\!\!\to} X$, and where the second line follows from the first by linearity of expectation.

Remark 2.15. It is also often of interest to consider *parameterized* channels, for which we can use the **Para** construction [12]. This acts by adjoining an object of parameters to the domain of the category at hand (such as the category of Bayesian lenses), and tensoring parameters of composite morphisms. Formally, this corresponds to a generalization of the indexed category of state-dependent morphisms, and forms the subject of a parallel submission. The parameters might represent the 'weights' of a neural network, or encode some structure about the possible predictions. Lacking the space to do justice to this structure here, we nonetheless leave it at that.

3 Systems Within Interfaces; Worlds Within Worlds

In this section, we develop the structures required for extending the formalism of statistical games to embodied systems.

3.1 Polynomials for Embodiment and Interaction

Each system in our universe inhabits some interface or boundary. It receives signals from its environments through this boundary, and can act by changing its shape (and, as we will see later, its position). As a system changes its shape, the set of possible immanent signals might change accordingly: consider a hedgehog rolling itself into a ball, thereby protecting its soft underbelly from harm (amongst other immanent signals). A system may also change its shape by coupling itself to some other system, such as when we pick up chalk to work through a problem. And shapes can be abstract: we change our 'shapes' when we enter an online video conference, or move within a virtual reality. We describe all of these interactions formally using polynomial functors, drawing on the work of [13].

Definition 3.1. Let $\mathscr{E}$ be a locally Cartesian closed category, and denote by y^A the representable copresheaf $y^A := \mathscr{E}(A, -) : \mathscr{E} \to \mathscr{E}$. A *polynomial functor* p is a coproduct of representable functors, written $p := \sum_{i:p(1)} y^{p_i}$, where $p(1) : \mathscr{E}$ is the indexing object. The category of polynomial functors in $\mathscr{E}$ is the full subcategory $\mathbf{Poly}_{\mathscr{E}} \hookrightarrow [\mathscr{E}, \mathscr{E}]$ of the $\mathscr{E}$-copresheaf category spanned by coproducts of representables. A morphism of polynomials is therefore a natural transformation.

Remark 3.2. Every copresheaf $P : \mathscr{E} \to \mathscr{E}$ corresponds to a bundle $p : E \to B$ in $\mathscr{E}$, for which $B = P(1)$ and for each $i : P(1)$, the fibre p_i is $P(i)$. We will henceforth elide the distinction between a copresheaf P and its corresponding bundle p, writing $p(1) := B$ and $p[i] := p_i$, where $E = \sum_i p[i]$. A natural transformation $f : p \to q$ between copresheaves therefore corresponds to a map of bundles. In the case of polynomials, by the Yoneda lemma, this map is given by a 'forwards' map $f_1 : p(1) \to q(1)$ and a family of 'backwards' maps $f^\# : q[f_1(\text{-})] \to p[\text{-}]$ indexed by $p(1)$, as in the left diagram below. Given $f : p \to q$ and $g : q \to r$, their composite $g \circ f : p \to r$ is as in the right diagram below.

$$
\begin{array}{ccccc}
E & \xleftarrow{f^\#} & f^*F & \longrightarrow & F \\
{\scriptstyle p}\downarrow & & \downarrow & \lrcorner & \downarrow{\scriptstyle q} \\
B & = & B & \xrightarrow{f_1} & C
\end{array}
\qquad\qquad
\begin{array}{ccccc}
E & \xleftarrow{(gf)^\#} & f^*g^*G & \longrightarrow & G \\
{\scriptstyle p}\downarrow & & \downarrow & \lrcorner & \downarrow{\scriptstyle r} \\
B & = & B & \xrightarrow{g_1 \circ f_1} & D
\end{array}
$$

where $(gf)^\#$ is given by the $p(1)$-indexed family of composite maps $r[g_1(f_1(\text{-}))] \xrightarrow{f^*g^\#} q[f_1(\text{-})] \xrightarrow{f^\#} p[\text{-}]$.

In our morphological semantics, we will call a polynomial p a *phenotype*, its base type $p(1)$ its *morphology* and the total space $\sum_i p[i]$ its *sensorium*. We will call elements of the morphology *shapes* or *configurations*, and elements of the sensorium *immanent signals*.

Proposition 3.3 ([13]). There is a monoidal structure $(\mathbf{Poly}_{\mathscr{E}}, \otimes, y)$ that we interpret as "putting systems in parallel". Given $p : \sum_i p[i] \to p(1)$ and $q : \sum_j q[j] \to q(1)$, we have $p \otimes q = \sum_i \sum_j p[i] \times q[j] \to p(1) \times q(1)$. $y : 1 \to 1$ is then clearly unital. □

Proposition 3.4 ([13]). The monoidal structure $(\mathbf{Poly}_{\mathscr{E}}, \otimes, y)$ is closed, with corresponding internal hom denoted $[-,-]$.

We interpret morphisms $(f_1, f^\#)$ of polynomials as encoding interaction patterns; in particular, such morphisms encode how composite systems act as unities. For example, a morphism $f : p \otimes q \to r$ specifies how the systems p and q come together to form a system r: the map f_1 encodes how r-configurations are constructed from configurations of p and q; and the map $f^\#$ encodes how immanent signals on p and q result from signals on r or from the interaction of p and q. For intuition, consider two people engaging in a handshake, or an enzyme acting on a protein to form a complex. The internal hom $[o, p]$ encodes all the possible ways that an o-phenotype system can "plug into" a p-phenotype system.

Remark 3.5. In the literature on active inference and the free energy principle, there is much debate about the concept of 'Markov blanket', an informal notion conceived to represent the boundary of an adaptive system. We believe that the algebra of polynomials is sufficient to formalize this concept precisely, and clear up much of the confusion in the literature.

3.2 Dynamical Behaviours on Polynomial Interfaces

Although dynamical systems can be modelled within $\mathbf{Poly}_{\mathscr{E}}$ itself, we prefer to take a fibrational perspective, following the idea that polynomials represent the boundaries of systems, separating 'internal' states from 'external'. We will therefore adopt a pattern of indexing categories by polynomials: in the case of dynamics, the fibre over a polynomial will be a category of possible internal systems, whose projection forgets the internal structure. We adopt an observational perspective on dynamics accordingly: our categories of dynamics will in fact be categories of behaviours compatible with the phenotype. We follow [8] and use sheaves over a site of intervals, sections of which constitute finite trajectories. Adopting such a topos gives us access to its corresponding logic and temporal type theory [7], which in turn permits the consideration of statistical games in nonstationary contexts.

Definition 3.6. Let $(\mathbb{T},+,0)$ be a monoid representing time, such as $\mathbb{R}_+$ or $\mathbb{N}$. Define the **interval site** to be the twisted arrow category of the delooping of $\mathbb{T}$, denoted $\mathbf{Int}_{\mathbb{T}} := \mathbf{Tw}(\mathbf{B}\mathbb{T})$, equipped with the Johnstone coverage [8]. Explicitly, the objects of $\mathbf{Int}_{\mathbb{T}}$ are elements of $\mathbb{T}$ that we take to represent intervals, and the morphisms $a : l \to l'$ are inclusions of intervals l into l' at offset a, such that $a + l \le l'$.

Definition 3.7. Denote by $\mathscr{B}^{\mathbb{T}}_{\mathscr{E}}$ the category of $\mathscr{E}$-valued sheaves on $\mathbf{Int}_{\mathbb{T}}$.

Example 3.8. There is a functorial inclusion $\mathsf{Traj} : \mathscr{E} \hookrightarrow \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}$ of $\mathscr{E}$ into $\mathscr{B}^{\mathbb{T}}_{\mathscr{E}}$ given by $\mathsf{Traj}(X) = \mathscr{E}(-,X) : {\mathbf{Int}_{\mathbb{T}}}^{\mathrm{op}} \to \mathscr{E}$. Semantically, $\mathsf{Traj}(X)$ is the sheaf of trajectories over X.

Proposition 3.9. A dynamical system with state spaces in $\mathscr{E}$ is an object in the functor category $\mathbf{Cat}(\mathbf{B}\mathbb{T},\mathscr{E})$. There is an embedding $\mathbf{Cat}(\mathbf{B}\mathbb{T},\mathscr{E}) \hookrightarrow \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}$ given as follows. Suppose θ is a dynamical system on $X : \mathscr{E}$. The embedding takes θ to the subsheaf $\Theta \hookrightarrow \mathsf{Traj}(X)$ given by $\{\theta(-,x) \mid x : X\}$.

Definition 3.10. Define the $\mathbf{Poly}_{\mathscr{E}}$-indexed category $\mathbf{BP}^{\mathbb{T}}_{\mathscr{E}} : \mathbf{Poly}_{\mathscr{E}} \to \mathbf{Cat}$ of **dynamical behaviours over polynomial interfaces in $\mathscr{E}$ and time $\mathbb{T}$** as follows. On objects, let $\mathbf{BP}^{\mathbb{T}}_{\mathscr{E}}(p) := \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{i:p(1)} p[i]\right)$. On morphisms $\varphi : p \to q$, define $\mathbf{BP}^{\mathbb{T}}_{\mathscr{E}}(\varphi) : \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{i:p(1)} p[i]\right) \to \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{j:q(1)} q[j]\right)$ as follows. Let $\mathsf{Traj}(\varphi^{\#})^* : \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{i:p(1)} p[i]\right) \to \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\varphi^* \sum_{j:q(1)} q[j]\right)$ be the base-change along $\mathsf{Traj}(\varphi^{\#})$. Let $\pi_q : \varphi^* \sum_{j:q(1)} q[j] \to \sum_{j:q(1)} q[j]$ be the projection out of the pullback and let $\mathsf{Traj}(\pi_q)_! : \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\varphi^* \sum_{j:q(1)} q[j]\right) \to \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{j:q(1)} q[j]\right)$ be the corresponding left adjoint to base change along $\mathsf{Traj}(\pi_q)$. Then $\mathbf{BP}^{\mathbb{T}}_{\mathscr{E}}(\varphi) := \mathsf{Traj}(\pi_q)_! \circ \mathsf{Traj}(\varphi^{\#})^* : \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{i:p(1)} p[i]\right) \to \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(\sum_{j:q(1)} q[j]\right)$. When the choice of time monoid $\mathbb{T}$ is clear from the context, we will often write just $\mathbf{BP}_{\mathscr{E}}$.

Since we are principally interested in systems with uncertainty, we consider how to model *random* dynamical systems in this setting. Let $\mathscr{P} : \mathscr{E} \to \mathscr{E}$ be a probability monad on $\mathscr{E}$. The slice $1/\mathscr{K}\ell(\mathscr{P})$ of its Kleisli category under the terminal object 1 is the category of probability spaces with measure-preserving maps in $\mathscr{E}$. A random dynamical system is a bundle of dynamical systems in $\mathscr{E}$, where the base system is measure-preserving. Denote the base probability space by (Ω,ν). The behaviours of a random dynamical system system then form a sheaf in $\mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}(\Omega)$ where the projection of the stalk at any point in time onto Ω is isomorphic to the support of ν.

Definition 3.11. We first define a polynomially-indexed category of **bundles of behaviours over probability spaces**. Let $U : 1/\mathscr{K}\ell(\mathscr{P}) \to \mathscr{E}$ be the forgetful functor taking a probability space (Ω,ν) to the corresponding space Ω in $\mathscr{E}$. Then define $\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}} : \mathbf{Poly}_{\mathscr{E}} \to \mathbf{Cat}$ on each p to be the lax colimit

$$\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(p) := \int^{(\Omega,\nu):1/\mathscr{K}\ell(\mathscr{P})} \mathscr{B}^{\mathbb{T}}_{\mathscr{E}}/\mathsf{Traj}\left(U(\Omega,\nu) \times \sum_{i:p(1)} p[i]\right).$$

On morphisms of polynomials $\varphi : p \to q$ we define $\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(\varphi)$ as in Definition 3.10:

$$\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(\varphi) := \int^{(\Omega,\nu)} \mathsf{Traj}\left(U(\Omega,\nu) \times \pi_q\right)_! \circ \mathsf{Traj}\left(U(\Omega,\nu) \times \varphi^{\#}\right)^*$$

We define the indexed category of random behaviours as a subcategory of these bundle systems.

Definition 3.12. Define the polynomially indexed **category of random behaviours** $\mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}} : \mathbf{Poly}_{\mathscr{E}} \to \mathbf{Cat}$ to be, on polynomials p, the subcategory of $\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(p)$ of sheaves whose stalks, when projected onto Ω, are isomorphic to the support of the corresponding measure ν. On morphisms of polynomials $\varphi : p \to q$, $\mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}(\varphi)$ is defined as $\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(\varphi)$, which preserves the stalk projection property since $\mathbf{BBP}^{\mathbb{T}}_{\mathscr{P}}(\varphi)$ doesn't act on the Ω factor.

Example 3.13. The solutions $X(t,\omega;x_0) : \mathbb{R}_+ \times \Omega \times M \to M$ to a stochastic differential equation $\mathrm{d}X_t = f(t,X_t)\mathrm{d}t + \sigma(t,X_t)\mathrm{d}W_t$, where $W : \mathbb{R}_+ \times \Omega \to M$ is a Wiener process in M, define a random dynamical system $\mathbb{R}_+ \times \Omega \times M \to M : (t,\omega,x) \mapsto X(t,\omega;x_0)$ over the Wiener base flow $\vartheta : \mathbb{R}_+ \times \Omega \to \Omega : (t,\omega) \mapsto W(s+t,\omega) - W(t,\omega)$ for any $s : \mathbb{R}_+$. We can alternatively represent this system as a bundle system over Ω, writing $\Theta : \mathbb{R}_+ \times \Omega \times M \to \Omega \times M : (t,\omega,x) \mapsto (\vartheta(t,\omega), X(t,\omega;x_0))$. We then seek a bundle of sheaves $S \to \mathsf{Traj}(\Omega)$. First, let $\bar{\Omega} := \{\vartheta(-,\omega) : \mathbb{R}_+ \to \Omega\}$ with ϑ being the Wiener base flow as given. Let $S := \{\Theta(-,\omega,x) : \mathbb{R}_+ \to \Omega \times M\}$. Then we have $S \to \bar{\Omega} \hookrightarrow \mathsf{Traj}(\Omega)$ as required.

Example 3.14. A Markov chain on M is given by a coalgebra $\tau : M \to \mathscr{P}M$. This is equivalently a pair $(\Omega \times M \xrightarrow{\tau^\#} M, \nu : \mathscr{P}\Omega)$. So now suppose $\vartheta : \mathbb{N} \times \Omega \to \Omega$ is a ν-preserving system. Given a sequence of points $\{\omega_i\}_{i:\mathbb{N}}$ we can form the system $\tau' : \mathbb{N} \times M \to M : (n,x) \mapsto \tau^\#(\omega_n) \circ \cdots \circ \tau^\#(\omega_0, x)$. But $\vartheta : \mathbb{N} \times \Omega \to \Omega$ generates such a sequence in Ω, so we can define $\bar{\tau} : \mathbb{N} \times \Omega \times M \to M : (n,\omega,x) \mapsto \tau^\#\left(\vartheta^{n-1}(\omega)\right) \circ \cdots \circ \tau^\#(\omega,x)$ when $n > 0$ and $\bar{\tau}(0,\omega,x) = x$, which satisfies the cocycle condition of a (random) dynamical system. We then obtain a trivial bundle system on $\Omega \times M \to \Omega$ by defining $\Theta : \mathbb{N} \times \Omega \times M \to \Omega \times M : (n,\omega,x) \mapsto (\vartheta(n,\omega), \bar{\tau}(n,\omega,x))$ and then obtain a bundle of behaviour sheaves by the method in Example 3.13.

3.3 Nested Systems and Dependent Polynomials

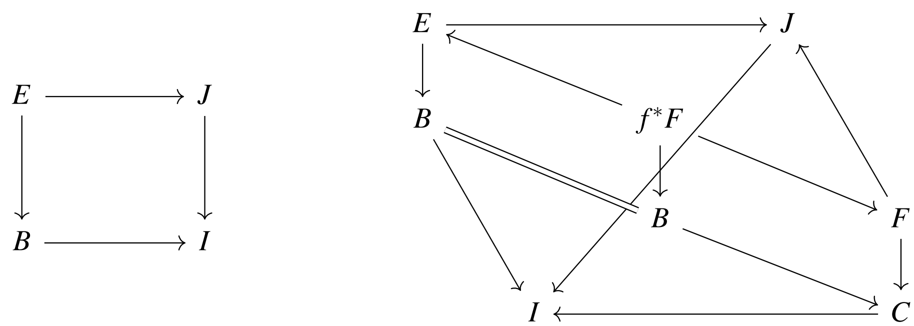
The foregoing formalism suffices to describe systems' shapes, and behaviours of those shapes that depend on their sensoria. But in our world, a system has a *position* as well as a shape! Indeed, one might want to consider systems nested within systems, such that the outer systems constitute the 'universes' of the inner systems; in this way, inner shapes may depend on outer shapes, and inner sensoria on outer sensoria.[2] We can model this situation polynomially.

Recall that an object in $\mathbf{Poly}_{\mathscr{E}}$ corresponds to a bundle $E \to B$, equivalently a diagram $1 \leftarrow E \xrightarrow{p} B \to 1$, and note that the unit polynomial y corresponds to a bundle $1 \to 1$. We can then think of $\mathbf{Poly}_{\mathscr{E}}$ as the category of "polynomials in one variable", or "polynomials over y". This presents a natural generalization, to polynomials in many variables, corresponding to diagrams $J \leftarrow E \to B \to I$; these diagrams form the objects of a category $\mathbf{Poly}_{\mathscr{E}}(J,I)$. When J is a (polynomial) bundle β over I, then we can take the subcategory of $\mathbf{Poly}_{\mathscr{E}}(J,I)$ whose objects are commuting squares and whose morphisms are prisms as follows; the commutativity ensures that inner and outer sensoria are compatible.

Proposition 3.15. There is an indexed category of **nested polynomials** which by abuse of notation we will call $\mathbf{Poly}_{\mathscr{E}}(-) : \mathbf{Poly}_{\mathscr{E}} \to \mathbf{Cat}$. Given $\beta : J \to I$, the category $\mathbf{Poly}_{\mathscr{E}}(\beta)$ has commuting squares as on the left below as objects and prisms as on the right as morphisms. Its action on polynomial morphisms

[2] We might even consider the outer shapes explicitly as positions in some world-space, and the outer sensorium as determined by possible paths between positions, in agreement with the perspective of [13] on polynomials.

$\beta \to \gamma$ is given by composition.



Remark 3.16. This construction can be repeatedly iterated, modelling systems within systems within systems. We leave the consideration of the structure of this iteration to future work, though we expect it to have an opetopic shape equivalent to that obtained by iterating the **Para** construction (Remark 2.15).

Observation 3.17. Our polynomially indexed categories of dynamical behaviours (Def. 3.10) and statistical games (Prop. 4.6) generalize to the case of nested polynomials, giving a doubly-indexed structure.

4 Theories of Approximate and Active Inference

We now start to bring together the structures of the previous sections, in order to breathe life into polynomials. We begin by sketching *approximate inference doctrines*, which characterize dynamical systems that optimize their performance at statistical games, without reference to morphology. In this paper, we do not concentrate on the detailed structure of these doctrines, leaving their exposition and comparison to future work, where we will also be interested in morphisms between doctrines.

4.1 Two Approximate Inference Doctrines

An approximate inference doctrine will be a monoidal functor from a category of statistical games into an appropriate category of dynamical systems, taking games to systems that 'play' those games, typically by implementing an optimization process. In the free-energy literature (*e.g.*, [2]), these systems have a hierarchical structure in which the realization of a game has access to the dynamics realizing the prior, mirroring the context-dependence of the games themselves. We therefore define the dynamical semantics accordingly.

Proposition 4.1. Let $\mathscr{P} : \mathscr{E} \to \mathscr{E}$ be a probability monad and $\mathbb{T}$ be a time monoid. There is a monoidal category of **Bayesian lenses with dynamical semantics**, denoted $\mathbf{DynLens}^{\mathbb{T}}_{\mathscr{P}}$, whose objects are those of $\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}$ and whose morphisms $(X,A) \nrightarrow (Y,B)$ are lenses equipped with dynamical semantics functors $\mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}(Xy^A) \to \mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}(AYy^{XB})$. Given lenses $(X,A) \nrightarrow (Y,B) \nrightarrow (Z,C)$, we form the composite semantics functor $\mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}(Xy^A) \to \mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}(AZ^{XC})$ by first obtaining a system on AYy^{XB} using the semantics of the first factor, wiring the Xy^A system to obtain a system on AYy^{AB}, tracing out the A to obtain Yy^B, and hence BZ^{YC} from the second semantics functor; we then tensor AYy^{XB} and BZ^{YC}, obtaining $ABYZy^{XYBC}$, tracing out BY to obtain AZy^{XC} as required. The monoidal product on semantics functors is given by tensoring the systems.

Definition 4.2. An **approximate inference doctrine** is a monoidal functor from $\mathbf{SGame}_{\mathscr{C}}$ (or some subcategory thereof) for some Markov category $\mathscr{C}$ to $\mathbf{DynLens}^{\mathbb{T}}_{\mathscr{P}}$ for some probability monad $\mathscr{P}$.

Many informal approximate inference schemes—including Markov chain Monte Carlo, variational Bayes, expectation-maximization, particle filtering—give rise to approximate inference doctrines; functoriality typically follows from Theorem 2.5. Here we note two explicitly, for later reference.

Proposition 4.3. Let $\mathscr{C}$ be the subcategory of $\mathscr{K}\ell\mathscr{P}$ spanned by Euclidean spaces and stochastic channels emitting Gaussian distributions. Let $\mathscr{G}$ be the subcategory of $\mathbf{SGame}_{\mathscr{K}\ell(\mathscr{P})}$ restricted to free energy games. The *Laplace doctrine* takes each such game to a dynamical system performing gradient descent on the free energy under the assumption that the Gaussians are tightly concentrated about their modes such that the free energy is well approximated by a Taylor expansion to 2nd order.[2] (Under this assumption, analytic expressions for the dynamics obtain.)

Proposition 4.4. By lifting the category $\mathscr{C}$ of the preceding proposition to the behaviour topos $\mathscr{B}_{\mathscr{E}}$, we can define a category of "free action" games: these are games whose loss functions correspond to the time-integral of the free-energy over the trajectories. By making the same assumptions of tightly-peaked Gaussians, one obtains the approximate inference doctrine of generalized filtering [6]; the resulting equations of motion are formally similar, but now encompass nonstationary contexts.

4.2 Statistical Games over Polynomials

Despite admitting dynamical contexts, the approximate inference doctrine of generalized filtering nonetheless does not supply a satisfactory model of active systems. One piece of structure is still missing, with which we can describe action and interaction faithfully: an indexed category of statistical games over polynomials. In order to construct this, we first define categories of "games on interfaces": this is simpler than slicing the category of statistical games, as we do not require games between games.

Definition 4.5. Let $\mathscr{P} : \mathscr{E} \to \mathscr{E}$ be a probability monad on $\mathscr{E}$. Let $X : \mathscr{E}$ be an object in $\mathscr{E}$. Define a category of **simple statistical games on the interface** X, denoted $\mathbf{IntGame}_{\mathscr{P}}(X)$, as follows. Its objects are simple statistical games with codomain X; that is, points of $\sum_{A:\mathscr{E}} \mathbf{SimpSGame}_{\mathscr{K}\ell(\mathscr{P})}(A,X)$. Let $(\gamma,\rho,\phi) : A \to X$ and $(\delta,\sigma,\chi) : B \to X$ be two such simple statistical games. Then a morphism $(\gamma,\rho,\phi) \to (\delta,\sigma,\chi)$ is a deterministic function $f : A \to B$—that is, a point of $\mathscr{E}(A,B)$—such that $\gamma = \delta \circ f$. Unitality and associativity follow immediately from those properties in $\mathscr{E}/\mathscr{P}X$.

We then use this to construct games over polynomials. The intuition here is that 'inside' a system with a polynomial phenotype is a statistical model of the system's sensorium. This involves an object representing the space of possible causes of observations, and a simple statistical game from this object onto the sensorium; by its nature, this model induces predictions about the system's configurations, as well about the immanent signals. By sampling from these predicted configurations, the system can act; by observing its actual configuration and the corresponding immanent signals, it can update its internal beliefs, and any parameters of the model. Later, we will equip this process with (random) dynamics, thereby giving the systems life.

Proposition 4.6. Let $\mathscr{P} : \mathscr{E} \to \mathscr{E}$ be a probability monad on a locally Cartesian closed category $\mathscr{E}$. There is a **polynomially indexed category of statistical games** $\mathbf{PSGame}_{\mathscr{P}} : \mathbf{Poly}_{\mathscr{E}} \to \mathbf{Cat}$, defined on objects p as $\mathbf{IntGame}_{\mathscr{P}}\left(\sum_{i:p(1)} p[i]\right)$. We defer the definition of the action of $\mathbf{PSGame}_{\mathscr{P}}$ on polynomial maps to §A.2, where we also give the proof of pseudofunctoriality.

Example 4.7. To understand the action of $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ on statistical games, it may help to consider the example of a corporation. Such a system is composed of a number of active systems, instantiating statistical games, interacting according to some pattern, formalized by the polynomial map ϕ. Given such a collection of games $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ tells us how to construct a game for the corporation as a whole: in particular, we obtain a stochastic channel generating predictions for the (exposed) sensorium of the corporation, and an inversion updating the constituent systems' beliefs accordingly.

4.3 Active Inference

We are now ready to define active inference doctrines; given all the foregoing structure, this proves relatively simple.

Definition 4.8. Let $\mathscr{P} : \mathscr{E} \to \mathscr{E}$ be a probablity (valuation) monad on a locally Cartesian closed category, and let $\mathbb{T}$ be a time monoid. An **active inference doctrine** is a monoidal indexed functor from $\mathbf{PSGame}_{\mathscr{P}}$ (or some sub-indexed category thereof) to $\mathbf{RBP}^{\mathbb{T}}_{\mathscr{P}}$.

Proposition 4.9. Both the Laplace and generalized filtering doctrines lift from approximate to active inference. The functors on each fibre are as before on games (here, objects), and on morphisms between games they are given merely by lifting the corresponding maps to the behaviour topos. One then checks that morphisms of polynomials correspond to natural transformations between these functors.

Remark 4.10. It would be desirable to incorporate the compositional structure of the games themselves, rather than treat them opaquely as objects. This suggests a double-categorical structure the investigation of which we leave to future work. Similarly, we do not here elaborate the extension of these indexed categories to the dependent-polynomial case.

5 Polynomial Life and Embodied Cognition

Finally, we sketch how a number of classic biological processes can be modelled as processes of active inference over polynomials. The key insight is that, by fixing the prior of an 'active' free-energy game to encode high-precision (low-variance) beliefs about the external state, we can induce the system to prefer acting (to reify those beliefs) over perceiving (*i.e.*, updating the beliefs to match perceptions). In doing so, one can induce *volition* or *goal-directedness* in the system. A key feature of these examples is that they demonstrate 'embodied' cognition, in which a system's form and interactions become part of its cognitive apparatus.

Remark 5.1. Of course, one must be careful not to choose a prior with excessively high precision (such as a Dirac delta distribution), as this would cause the system to forego any belief-updating, thereby rendering its actions independent from the 'actual' external state.

Example 5.2. Suppose that the system's sensorium includes a key parameter such as ambient temperature or blood pH. Suppose that by adjusting its configuration, the system can move around in order to sample this parameter. And suppose that the 'prior' encodes a high-precision distribution centred on the acceptable range of this parameter. Then it is straightforward to show that the system, by minimizing the free energy, will attempt to configure itself so as to remain within the acceptable parameter range. We can consider this as a simple model of **homeostasis**.

Example 5.3. We can extend the previous example to a system with multiple (polynomial) components, each equipped with a "homeostasis game", in order to model **morphogenesis**. Suppose the environmental parameter in the sensorium is the local concentration of some signalling molecule, and suppose the polynomial morphism forming the composite system encodes the pattern of signalling molecule concentrations in the neighbourhood of each system, as a result of their mutual configurations. Suppose then that the target state encoded in the prior of each system corresponds to the system being positioned in a particular way relative to the systems around it, as represented by the signal concentrations. Free-energy minimization then induces the systems to arrange themselves in order to obtain the target pattern.

Remark 5.4. The foregoing examples begin to point towards a compositional theory of *autopoiesis*: here, one might expect the target state to encode the proposition "maintain my morphology", which appears self-referential. The most elegant way of encoding this proposition in the prior is not immediately clear, although a number of possibilities present themselves (such as avoiding some undesirable configuration representing dissolution). We expect a satisfactory answer to this to be related to "Bayesian mechanics" (§6).

Remark 5.5. It has been shown informally that, given a finite time horizon Markov decision problem, active inference can recover the Bellman-optimal policy traditionally obtained by backward induction [4]. On-going work by the present author is directed at formalizing the structure of this relationship. In particular, the result rests on encoding directly into the prior the expectation of the loss function given a policy and a goal, which strikes us as a large amount of information to push into an unstructured distribution over numbers.

Example 5.6. The examples need not be restricted to simple biological cases. For instance, we can model spatial navigation quite generally: we can use parameterized statistical games to encode uncertainty about the structure of the 'external space' (for instance: which points or neighbourhoods are connected to which, and by which paths). By setting a high-precision prior at some location, the system will attempt to reach that location, learning the spatial structure along the way; reducing the precision of the prior causes the system to prefer "mere exploration". One can attach sense-data to each location using the natural polynomial bundle structure. Moreover, the 'external space' need not be a simple topological space: it may be something more structured. For instance, categories and sites can themselves be modelled polynomially. One can think of "taking an action" as precisely analogous to "following a morphism": thus, in a topos-theoretic setting, one can consider the structure of the 'external space' to be a type-theoretic context, and positions in the world to be objects in the corresponding topos. One could then encode in the prior a target proposition, and free-energy minimization would cause the system then to explore the 'space' (learning its structure), and seek a path to the target. But such a path is precisely a proof! There is increasing evidence that the neural mechanisms underlying spatial and abstract navigation are the same [1], and this seems to supply a mathematical justification.

6 Future Directions

Besides expanding the examples above in detail, there are many future directions to pursue. Our last example points towards a 'well-typed' theory of cognition, finding type-theoretic analogues of cognitive processes (such as action, planning, or navigation). By formalizing the connection between polynomial statistical games and Markov decision processes, we hope finally to relate our 'statistical' account of cybernetics with the account emerging from research in compositional game theory. In particular, we believe that the hierarchical/nested structure of our polynomial systems is structurally similar to that of (parameterized) players in open games. Along similar lines, we expect a connection between statistical games and 'learners' [12] implementing backprop.

In a more physical direction, there is a controversy in the informal literature about whether one should expect *any* system with a boundary (and hence any system on a polynomial) to admit a canonical statistical-game description; the typical suggestion is that such a description should obtain at non-equilibrium steady state, through a manipulation of the corresponding Fokker-Planck equation; this is the notion of "Bayesian mechanics" [5]. Our results suggest that such a canonical description should form a left adjoint to some active inference doctrine; this is a matter of on-going research by the author.

Working topos-theoretically points further in a metaphysical direction: a Bayesian perspective lends itself to subjectivism, but considering the "internal universe" of a navigating system to be a topos in some context also points to a subjective realism. It seems likely then that composite systems need not in general agree about their observations. We should therefore expect to find evidence of contextuality and disagreement in multi-agent systems, and to investigate this using cohomological tools (*e.g.*, [3]).

References

[1] Timothy EJ Behrens, Timothy H Muller, James CR Whittington, Shirley Mark, Alon B Baram, Kimberly L Stachenfeld & Zeb Kurth-Nelson (2018): *What is a cognitive map? Organizing knowledge for flexible behavior*. Neuron 100(2), pp. 490–509.

[2] Christopher L Buckley, Chang Sub Kim, Simon McGregor & Anil K Seth (2017): *The free energy principle for action and perception: A mathematical review*. Journal of Mathematical Psychology 81, pp. 55–79.

[3] Giovanni Carù (2017): *On the Cohomology of Contextuality*. *EPTCS 236, 2017, pp. 21-39*, doi:10.4204/EPTCS.236.2.

[4] Lancelot Da Costa, Noor Sajid, Thomas Parr, Karl Friston & Ryan Smith (2020): *The Relationship between Dynamic Programming and Active Inference: The Discrete, Finite-horizon Case*.

[5] Karl Friston (2019): *A free energy principle for a particular physics*.

[6] Karl Friston, Klaas Stephan, Baojuan Li & Jean Daunizeau (2010): *Generalised Filtering*. Mathematical Problems in Engineering 2010, pp. 1–34, doi:10.1155/2010/621670.

[7] Patrick Schultz & David I. Spivak: *Temporal Type Theory: A topos-theoretic approach to systems and behavior*. Springer, Berkhauser.

[8] Patrick Schultz, David I Spivak & Christina Vasilakopoulou (2019): *Dynamical Systems and Sheaves*, pp. 1–57.

[9] Toby St. Clere Smithe: *Bayesian Updates Compose Optically*.

[10] Toby St Clere Smithe: *Cyber Kittens, or Some First Steps Towards Categorical Cybernetics*. In: Proceedings 3rd Annual International Applied Category Theory Conference 2020 (ACT 2020).

[11] David I. Spivak: *Generalized Lens Categories via functors $\mathscr{C}^{op} \to$ **Cat***.

[12] David I. Spivak: *Learners' languages*.

[13] David I. Spivak: *Poly: An abundant categorical setting for mode-dependent dynamics*.

A Proofs

A.1 Proof of Prop 2.9

Proof. Suppose given statistical games $(f,\phi) : (X,A) \to (Y,B)$ and $(g,\psi) : (Y,B) \to (Z,C)$. We seek a composite game $(g,\psi) \circ (f,\phi) := (gf, \psi\phi) : (X,A) \to (Z,C)$. We have $gf = g \,\mathbin{⚬}\, f$ by lens composition. Propositon 2.8 gives us a dependent function $\mathsf{localCtx}$ with signature

$$\overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Z,C)\big) \times \mathbf{BayesLens}_{\mathscr{C}}\big((X,A),(Y,B)\big) \times \mathbf{BayesLens}_{\mathscr{C}}\big((Y,B),(Z,C)\big)$$
$$\to \overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((X,A),(Y,B)\big) \times \overline{\mathbf{BayesLens}_{\mathscr{C}}}\big((Y,B),(Z,C)\big) .$$

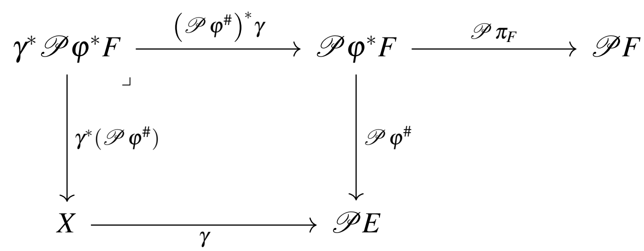
We therefore define the composite fitness function $\psi\phi$ to be

$$\psi\phi := \mathsf{add} \circ (\phi, \psi) \circ \mathsf{localCtx}(-, f, g)$$

where $\mathsf{add} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the addition of real numbers. (Note that in general any associative unital binary operation on $\mathbb{R}$ would suffice here.) The identity game $(X, A) \to (X, A)$ is given by $(\mathsf{id}, 0)$, the pairing of the identity lens on (X, A) with the constant function on 0 (or, more generally, constant on the unit of the algebra add). Associativity and unitality are immediate from associativity and unitality of lens composition and associativity and unitality of add. □

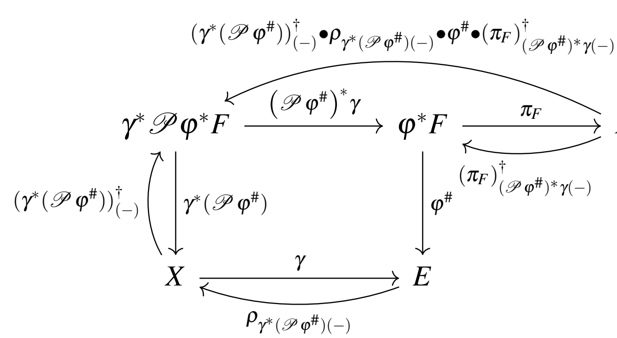
A.2 Details for Proposition 4.6

Proposition A.1 (Action of $\mathbf{PSGame}_{\mathscr{P}}$ on morphisms of polynomials)**.** A morphism of polynomials $\varphi : p \to q$ induces a functor $\mathbf{PSGame}_{\mathscr{P}}(\varphi) : \mathbf{PSGame}_{\mathscr{P}}(p) \to \mathbf{PSGame}_{\mathscr{P}}(q)$. Suppose the polynomial p corresponds to a bundle $E \to B$, where $E = \sum_{i:p(1)} p[i]$ and $B = p(1)$, and q corresponds to a bundle $F \to C$ with $F = \sum_{j:q(1)} q[j]$ and $C = q(1)$. Write $(\gamma, \rho, \phi) : X \to E$ for a generic statistical game on p: this consists in a forwards channel $\gamma : X \to \mathscr{P}E$, an inverse channel $\rho : \mathscr{P}X \times E \to \mathscr{P}A$, and a fitness function $\phi : \overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}((X,X),(E,E)) \to \mathbb{R}$. We now define the action of $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ on each of these three components. On the forwards channel, $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ takes γ to the composite map in the top of the following diagram in $\mathscr{E}$, where the square defines the indicated pullback:



(5)

On the inverse channel, $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ takes ρ to the composite map indicated at the top of the following diagram in $\mathscr{K}\ell(\mathscr{P})$. Note that, for notational simplicity, we have used the name of a map in $\mathscr{E}$ for its image under the canonical lifting by Kleisli extension to $\mathscr{K}\ell(\mathscr{P})$.



(6)

On fitness functions, $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ takes $\phi : \overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}\big((X,X),(E,E)\big) \to \mathbb{R}$ to a fitness function of type $\overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}\big((\gamma^* \mathscr{P}\varphi^* F, \gamma^* \mathscr{P}\varphi^* F),(F,F)\big) \to \mathbb{R}$, as follows. We can construct simple Bayesian lenses $\gamma^* \mathscr{P}\varphi^* F \nrightarrow X$ and $F \nrightarrow E$ using maps from the diagrams above. First, observe that $(\xi, \xi^\dagger) := \Big(\gamma^*(\mathscr{P}\varphi^\#), \big(\gamma^*(\mathscr{P}\varphi^\#)\big)^\dagger\Big)$ is a Bayesian lens of the former type. Next, observe that the first factor of the context gives a state ν on $\gamma^* \mathscr{P} \varphi^* F$, and pushing this through $\left(\mathscr{P} \varphi^\#\right)^* \gamma$ gives a state $\left(\mathscr{P} \varphi^\#\right)^* \gamma \bullet \nu$ on $\varphi^* F$. We now construct a simple Bayesian lens $(\zeta, \zeta^\dagger) : F \nrightarrow E$, writing

$$\zeta := F \xrightarrow{(\pi_F)^\dagger_{(\mathscr{P}\varphi^\#)^*\gamma}} \mathscr{P} \varphi^* F \xrightarrow{\mathscr{P} \varphi^\#} \mathscr{P} E \,,$$
$$\zeta^\dagger := \mathscr{P} F \times E \xrightarrow{(\pi_F)^\dagger_{(\mathscr{P}\varphi^\#)^*\gamma \bullet \nu} \times \mathsf{id}_E} \mathscr{P} \varphi^* F \times E \xrightarrow{(\varphi^\#)^\dagger} \mathscr{P} \varphi^* F \xrightarrow{\mathscr{P} \pi_F} \mathscr{P} F \,.$$

Hence we obtain a map

$$\beta := \overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}\big((\gamma^* \mathscr{P} \varphi^* F, \gamma^* \mathscr{P} \varphi^* F), (F, F)\big)$$
$$\xrightarrow{\overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}\left((\xi,\xi^\dagger),(\zeta,\zeta^\dagger)\right)} \overline{\mathbf{BayesLens}_{\mathscr{K}\ell(\mathscr{P})}}\big((X, X), (E, E)\big)$$

by post- and pre-composition. Then composing $\phi \circ \beta$ gives a fitness function of the required type.

Next, we define the action of $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ on morphisms $f : X \to Y$ in $\mathbf{PSGame}_{\mathscr{P}}(p)$. Let $(\gamma, \rho, \phi) : X \to E$ and $(\delta, \sigma, \chi) : Y \to E$ be the objects (statistical games with codomains E). Note that $\gamma^* \mathscr{P} \varphi^* F \cong (\mathscr{P} \varphi^\#)^* \gamma$ and $\delta^* \mathscr{P} \varphi^* F \cong (\mathscr{P} \varphi^\#)^* \delta$. Then let $\mathbf{PSGame}_{\mathscr{P}}(\varphi)(f) := (\mathscr{P} \varphi^\#)^* f$. Commutativity of the corresponding triangle in $\mathbf{PSGame}_{\mathscr{P}}(q)$ is immediate, and functoriality follows from that of limits.

Proof of pseudofunctoriality. We treat $\mathbf{Poly}_{\mathscr{E}}$ as a trivial bicategory, with discrete hom-categories, and so it only remains to check pseudofunctoriality of $\mathbf{PSGame}_{\mathscr{P}}$; that is, that given $\varphi : p \to q$ and $\psi : q \to r$ in $\mathbf{Poly}_{\mathscr{E}}$, $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \varphi \cong \mathbf{PSGame}_{\mathscr{P}}(\psi) \circ \mathbf{PSGame}_{\mathscr{P}}(\varphi)$. We need to check this on forwards channels, backwards channels, fitness functions, and morphisms between games. Let the polynomial r correspond to the bundle $G \to D$ where $G := \sum_{k:r(1)} r[k]$ and $D := r(1)$.

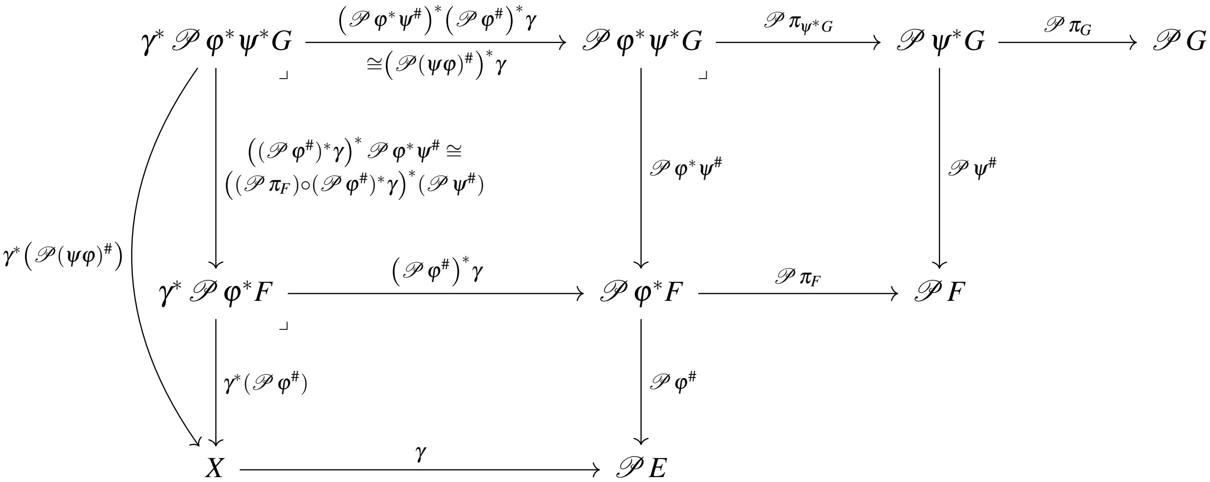
We start with forwards channels. The forwards channel of $\mathbf{PSGame}_{\mathscr{P}}(\psi) \circ \mathbf{PSGame}_{\mathscr{P}}(\varphi)(\gamma, \rho, \phi)$ is by definition the top map in the following diagram, where the bottom map below is the top map in the defining diagram (5) above:

$$\begin{array}{ccccc}
\gamma^* \mathscr{P} \varphi^* \psi^* G & \xrightarrow{\qquad\qquad\qquad} & \mathscr{P} \psi^* G & \longrightarrow & \mathscr{P} G \\
\big\downarrow & \lrcorner & \big\downarrow {\scriptstyle \mathscr{P} \psi^\#} & & \\
\gamma^* \mathscr{P} \varphi^* F \xrightarrow{\left(\mathscr{P} \varphi^\#\right)^* \gamma} \mathscr{P} \varphi^* F & \xrightarrow{\mathscr{P} \pi_F} & \mathscr{P} F & &
\end{array} \tag{7}$$

The forwards channel of $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \varphi)(\gamma, \rho, \phi)$ is by definition the top map in the following diagram:

$$\begin{array}{ccccc}
\gamma^* \mathscr{P} \varphi^* \psi^* G & \longrightarrow & \mathscr{P} \varphi^* \psi^* G & \longrightarrow & \mathscr{P} G \\
{\scriptstyle \gamma^*\left(\mathscr{P}(\psi\varphi)^\#\right)} \big\downarrow & \lrcorner & \big\downarrow {\scriptstyle \mathscr{P}(\psi\varphi)^\#} & & \\
X & \xrightarrow[\gamma]{} & \mathscr{P} E & &
\end{array} \tag{8}$$

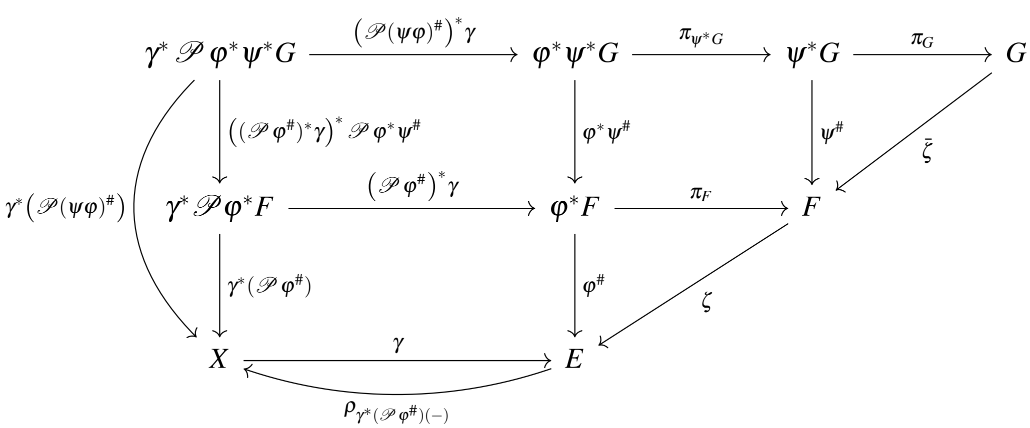
We need to show that the two composite top channels $\gamma^* \mathscr{P}\,\varphi^*\psi^* G \to \mathscr{P}\,G$ are equal up to isomorphism. Consider therefore the following commutative diagram, in which each square is a pullback.



(9)

The projection $\gamma^*\left(\mathscr{P}(\psi\varphi)^\#\right) : \gamma^*\,\mathscr{P}\,\varphi^*\psi^* G \to X$ (the left vertical map in (8)) factors as $\gamma^*\,\mathscr{P}\,\varphi^*\psi^* G \to \gamma^*\,\mathscr{P}\,\varphi^* F \to X$ by the universal properties of the defining pullbacks. This latter composite is the left vertical composite in (9) above. Next, notice that $\mathscr{P}(\psi\varphi)^\# \cong \mathscr{P}\,\varphi^\# \circ \mathscr{P}\,\varphi^*\psi^\#$ by the definition of $(\psi\varphi)^\#$ and functoriality of $\mathscr{P}$, and so the entire pullback square in (8) factors as the pasting of the vertical pair of pullback squares in (9). Consequently, the pullback square in (7) factors as the pasting of the horizontal pair of pullback squares in (9). Therefore, the top map in (7) is isomorphic to the top map in (9). Finally, note that the projection $\mathscr{P}\,\varphi^*\psi^* G \to \mathscr{P}\,G$ in (8) factors as $\mathscr{P}\,\varphi^*\psi^* G \to \mathscr{P}\,\psi^* G \to \mathscr{P}\,G$ by the universal properties of the defining pullbacks. Therefore, the top map in (8) is also isomorphic to the top map in (9). And so the top maps in (7) and (8) are isomorphic to each other, as required.

To check pseudofunctoriality on backwards channels and fitness functions, we lift diagram (9) to $\mathscr{K}\ell(\mathscr{P})$, as before, though now we include the maps ρ and ζ relevant for our present purposes. Note that this means that the diagram is not commutative, although each inner square commutes by the commutativity of the corresponding squares in (9).



(10)

Denote by δ by the action of $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ on γ, by σ the action of $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ on ρ, and by τ the action of $\mathbf{PSGame}_{\mathscr{P}}(\psi)$ on σ; that is,

$$\begin{aligned}
\delta &:= \mathscr{P}\,\pi_F \circ \left(\mathscr{P}\,\varphi^{\#}\right)^{*} \gamma, \\
\sigma &:= \left(\gamma^{*}(\mathscr{P}\,\varphi^{\#})\right)^{\dagger}_{(-)} \bullet \rho_{\gamma^{*}(\mathscr{P}\,\varphi^{\#})(-)} \bullet \varphi^{\#} \bullet (\pi_F)^{\dagger}_{(\mathscr{P}\,\varphi^{\#})^{*}\gamma(-)}, \\
\tau &:= \left(\delta^{*}(\mathscr{P}\,\psi^{\#})\right)^{\dagger}_{(-)} \bullet \sigma_{\delta^{*}(\mathscr{P}\,\psi^{\#})(-)} \bullet \psi^{\#} \bullet (\pi_G)^{\dagger}_{(\mathscr{P}\,\psi^{\#})^{*}\delta(-)}.
\end{aligned}$$

Denote by τ' the action of $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \varphi)$ on ρ; this is the backwards channel paired by $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \varphi)$ with the forwards channel at the top of diagram (10). We now check that τ' equals τ, reasoning diagrammatically. Note that (δ, σ) is a Bayesian lens. Let $\bar{\sigma}$ be defined as the factor of σ up to $(\pi_F)^{\dagger}$. Note then that $(\delta, \sigma) = \left(\pi_F, (\pi_F)^{\dagger}\right) \diamond \left(\left(\mathscr{P}\,\varphi^{\#}\right)^{*} \gamma, \bar{\sigma}\right)$, by Theorem 2.5. Reasoning similarly, τ factors through $(\pi_G)^{\dagger}$ and τ' factors through $(\pi_G \circ \pi_{\psi^{*}G})^{\dagger}$. Since the inner squares of (10) commute, any parallel path through them is equal. By Theorem 2.5, this also applies to the exact inversions. Furthermore, by following their paths through (10), we see that this applies specifically to τ and τ'. We therefore conclude that they are equal.

Next, we check pseudofunctoriality on fitness functions. $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ gives us, for a fitness function ϕ over p, two Bayesian lenses $(\xi, \xi^{\dagger})$ and $(\zeta, \zeta^{\dagger})$ which we compose with the lenses in the context over p, in order to obtain the fitness function $\phi \circ \beta$ over q. Denote by $(\bar{\xi}, \bar{\xi}^{\dagger})$ and $(\bar{\zeta}, \bar{\zeta}^{\dagger})$ the corresponding lenses obtained by applying $\mathbf{PSGame}_{\mathscr{P}}(\psi)$ to the fitness function $\phi \circ \beta$. We need to show that the lenses $\left(\xi', \xi'^{\dagger}\right)$ and $\left(\zeta', \zeta'^{\dagger}\right)$ obtained from applying $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \phi)$ to ϕ factor as $(\bar{\xi}, \bar{\xi}^{\dagger}) \diamond (\xi, \xi^{\dagger})$ and $(\zeta, \zeta^{\dagger}) \diamond (\bar{\zeta}, \bar{\zeta}^{\dagger})$. We have:

$$\begin{aligned}
(\xi, \xi^{\dagger}) &:= \left(\gamma^{*}(\mathscr{P}\,\varphi^{\#}), \left(\gamma^{*}(\mathscr{P}\,\varphi^{\#})\right)^{\dagger}\right), \\
(\bar{\xi}, \bar{\xi}^{\dagger}) &:= \left(\left((\mathscr{P}\,\varphi^{\#})^{*}\gamma\right)^{*} \mathscr{P}\,\varphi^{*}\psi^{\#}, \left(\left((\mathscr{P}\,\varphi^{\#})^{*}\gamma\right)^{*} \mathscr{P}\,\varphi^{*}\psi^{\#}\right)^{\dagger}\right), \\
(\xi', \xi'^{\dagger}) &:= \left(\gamma^{*}\left(\mathscr{P}(\psi\varphi)^{\#}\right), \left(\gamma^{*}\left(\mathscr{P}(\psi\varphi)^{\#}\right)\right)^{\dagger}\right).
\end{aligned}$$

Clearly by diagram (10) and Theorem 2.5, we have $\left(\xi', \xi'^{\dagger}\right) = (\bar{\xi}, \bar{\xi}^{\dagger}) \diamond (\xi, \xi^{\dagger})$. Similar "Bayesian diagram-chasing" demonstrates that $\left(\zeta', \zeta'^{\dagger}\right) = (\zeta, \zeta^{\dagger}) \diamond (\bar{\zeta}, \bar{\zeta}^{\dagger})$.

Finally, we check pseudofunctoriality on morphisms $f : X \to Y$. $\mathbf{PSGame}_{\mathscr{P}}(\varphi)$ takes f to $\left(\mathscr{P}\,\varphi^{\#}\right)^{*} f$; in turn, this is taken by $\mathbf{PSGame}_{\mathscr{P}}(\psi)$ to $\left(\mathscr{P}\,\psi^{\#}\right)^{*} \left(\mathscr{P}\,\varphi^{\#}\right)^{*} f$. Simultaneously, $\mathbf{PSGame}_{\mathscr{P}}(\psi \circ \varphi)$ takes f to $\left(\mathscr{P}(\psi\varphi)^{\#}\right)^{*} f$. The functoriality of limits immediately gives $\left(\mathscr{P}(\psi\varphi)^{\#}\right)^{*} f = \left(\mathscr{P}\,\psi^{\#}\right)^{*} \left(\mathscr{P}\,\varphi^{\#}\right)^{*} f$. □