Open dynamical systems as polynomial coalgebras

Toby St Clere Smithe

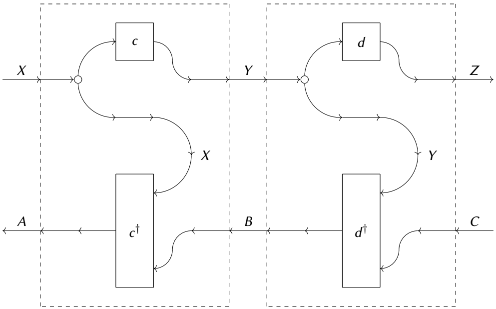
Topos Institute
&
University of Oxford

2022 July 19

Predictive coding: dynamical semantics for statistical games

I've been working on functors turning ...

Bayesian lenses:



into 'biological' dynamical systems:

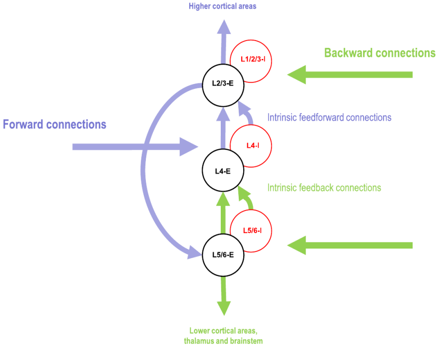


Figure: Bastos et al. [1]

This means finding categories of dynamics with the right compositional structure, which is (in some generality) what I'll be talking about today.

Outline of talk

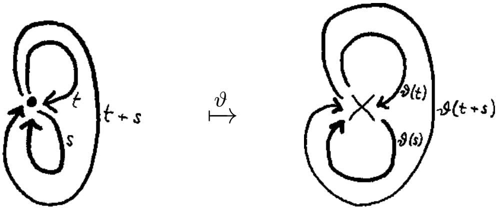
Deterministic closed dynamical systems

Given a monoid $\mathbb{T}$ and a space X, a *deterministic, closed* dynamical system is an action of $\mathbb{T}$ on X.

- *i.e.*, a family of maps $\{\vartheta(t) : X \to X\}_{t:\mathbb{T}}$ satisfying
  - $\vartheta(0) = \mathrm{id}_X$ "no evolution = no change"
  - $\vartheta(t) \circ \vartheta(s) = \vartheta(t+s)$ "evolving a bit then a bit more = evolving a little while"

Alternatively, if $\mathbf{B}\mathbb{T}$ is the delooping of $\mathbb{T}$ and X is an object of $\mathcal{E}$, then ϑ is a functor $\mathbf{B}\mathbb{T} \to \mathcal{E}$.

Deterministic closed dynamical systems are objects of the functor category $\mathbf{Cat}(\mathbf{B}\mathbb{T}, \mathcal{E})$.



(This perspective is, I believe, due to Lawvere.)

Examples of deterministic closed dynamical systems

1. In discrete time, with $\mathbb{T} = \mathbb{N}$, a system is simply a transition map: $\vartheta(t) = \vartheta(1)^{\circ t}$:
   - $\vartheta(0) = \mathrm{id}_X$; by induction: $\vartheta(t+1) = \vartheta(t) \circ \vartheta(1)$; $\vartheta(2) = \vartheta(1+1) = \vartheta(1) \circ \vartheta(1)$ *etc.*

2. Solutions to ordinary differential equations are systems with $\mathbb{T} = \mathbb{R}$:
   - $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x) \implies \vartheta(t) : x_0 \mapsto x(t)$ for each initial condition x_0

But what about other kinds of transition or system?

- For predictive coding, we need at least stochastic systems ...

Stochastic closed dynamical systems (and other transition types)

By changing the category $\mathcal{E}$, we obtain different types of transition.

In particular, we can choose the Kleisli category $\mathcal{K}\ell(M)$ of a monad $M : \mathcal{E} \to \mathcal{E}$.

- Same objects as $\mathcal{E}$; morphisms $\mathcal{K}\ell(M)(X, Y) = \{X \to MY\}$.
- This gives M-typed transitions: *e.g.*, stochastic when $M = \mathcal{D}$, the 'distribution' monad.

Example: A functor $\mathbf{B}\mathbb{T} \to \mathcal{K}\ell(\mathcal{D})$ is a *closed Markov semigroup* with time $\mathbb{T}$.

- Transitions satisfy $\vartheta_{s+t}(y|x) = \int_{x':X} \vartheta_s(y|x')\,\vartheta_t(\mathrm{d}x'|x)$
- Every (closed) Markov process induces a semigroup — *e.g.* SDE $\mathrm{d}X_t = f(X_t)\,\mathrm{d}W_t$

Note: a discrete time system $\vartheta : \mathbf{B}\mathbb{N} \to \mathcal{K}\ell(M)$ is fully specified by a *coalgebra* $X \to MX$

- Given an endofunctor $F : \mathcal{E} \to \mathcal{E}$, an M-coalgebra is just such a map

Polynomial functors: basic definition

We use polynomial functors to model the interfaces of our systems.

A *polynomial functor* is a coproduct of representable copresheaves.
Write $\mathcal{E}$ for the ambient category. (Ideally LCC, but we won't need all the structure today.)

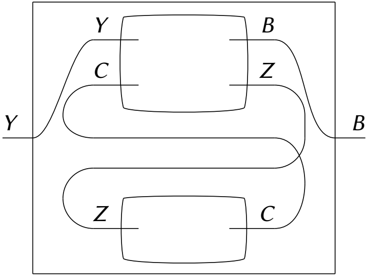
We follow Spivak in writing y^A for the representable copresheaf on A.
And we will write a general polynomial p as $\sum_{i:p(1)} y^{p[i]}$.

- $p(1)$ is obtained by applying p to the terminal object 1;
- each $p[i]$ is the representing object for the i^{th} summand.

Think of $p(1)$ as the set/space/type of possible configurations that a system can adopt.
For each $i : p(1)$, imagine $p[i]$ as the type of admissible inputs in configuration i.

Polynomial functors: tensor; morphisms

We can place systems side-by-side using the tensor $\otimes$.

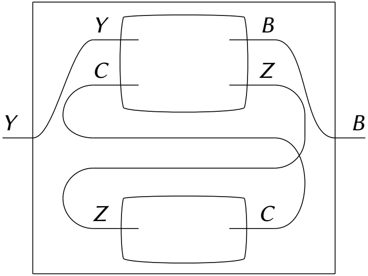


$$BZy^{YC} \otimes Cy^{Z} \to By^{Y}$$

And we can wire systems together using morphisms of polynomials.

Polynomial functors: tensor; morphisms

We can place systems side-by-side using the tensor $\otimes$.



$$BZy^{YC} \otimes Cy^{Z} \to By^{Y}$$

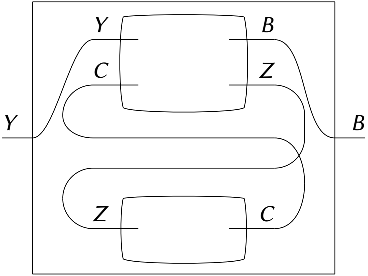
And we can wire systems together using morphisms of polynomials.

A little more precisely:

- A morphism $p \xrightarrow{\varphi} q$ of polynomials is a pair:
  - a forwards map $p(1) \xrightarrow{\varphi_1} q(1)$, and
  - a family of backwards maps $q[\varphi_1(i)] \xrightarrow{\varphi_i^\#} p[i]$.
- $\mathbf{Poly}_{\mathcal{E}}$ is equivalently the category of dependent lenses in $\mathcal{E}$: *i.e.*, the Grothendieck construction of the pointwise opposite of $\mathcal{E}/(-)$.

Polynomial functors: tensor; morphisms

We can place systems side-by-side using the tensor $\otimes$.



$$BZy^{YC} \otimes Cy^{Z} \to By^{Y}$$

And we can wire systems together using morphisms of polynomials.

A little more precisely:

- A morphism $p \xrightarrow{\varphi} q$ of polynomials is a pair:
  - a forwards map $p(1) \xrightarrow{\varphi_1} q(1)$, and
  - a family of backwards maps $q[\varphi_1(i)] \xrightarrow{\varphi_i^\#} p[i]$.
- **Poly**$_{\mathcal{E}}$ is equivalently the category of dependent lenses in $\mathcal{E}$: *i.e.*, the Grothendieck construction of the pointwise opposite of $\mathcal{E}/(-)$.
- The tensor $\otimes$ is given on objects by

$$p \otimes q := \sum_{i:p(1)} \sum_{j:q(1)} y^{p[i] \times q[j]}$$

and similarly on morphisms: take products in $\mathcal{E}$ of forwards and backwards maps.

Polynomial functors: internal hom; composition product

We need a couple more pieces of structure for our definition.

$(\otimes, y)$ has a corresponding internal hom $[-, =]$.

On objects: $[p, q] = \sum_{f : p \to q} y^{\sum_{i:p(1)} q[f_1(i)]}$

i.e., configurations are morphisms $p \xrightarrow{f} q$, and inputs are pairs (i, x) with i a p-configuration and x a corresponding q-input.

Polynomial functors: internal hom; composition product

We need a couple more pieces of structure for our definition.

$(\otimes, y)$ has a corresponding internal hom $[-, =]$.

On objects: $[p, q] = \sum_{f : p \to q} y^{\sum_{i:p(1)} q[f_1(i)]}$

i.e., configurations are morphisms $p \xrightarrow{f} q$, and inputs are pairs (i, x) with i a p-configuration and x a corresponding q-input.

Two useful facts:

1. $[Ay, y] \cong y^A$
2. morphisms $p \to y$ correspond to sections of the projection $\sum_{i:p(1)} p[i] \to p(1)$

Polynomial functors: internal hom; composition product

We need a couple more pieces of structure for our definition.

$(\otimes, y)$ has a corresponding internal hom $[-, =]$.

On objects: $[p, q] = \sum_{f : p \to q} y^{\sum_{i : p(1)} q[f_1(i)]}$

i.e., configurations are morphisms $p \xrightarrow{f} q$, and inputs are pairs (i, x) with i a p-configuration and x a corresponding q-input.

Two useful facts:

1. $[Ay, y] \cong y^A$
2. morphisms $p \to y$ correspond to sections of the projection $\sum_{i : p(1)} p[i] \to p(1)$

Composition of polynomials as functors induces a monoidal structure $(\triangleleft, y)$ by substitution.

Given polynomials p and q, $q \triangleleft p$ is the polynomial $q \circ p$ obtained by substituting $\sum_{i : p(1)} y^{p[i]}$ for each instance of y in q.

$\triangleleft$-comonoids are categories [2]: the counit encodes identities; the comultiplication encodes codomains and composites.

$\triangleleft$-comonoid homomorphisms are *cofunctors*[a].

[a] We'll hear more about these in the talk later today on enriched lenses by Clarke and Di Meglio.

For a fabulous exposition of the relevant details, see Spivak and Niu [3].

Deterministic open dynamical systems

Definition (2.1)

A deterministic open dynamical system with interface p, state space S and time $\mathbb{T}$ is a polynomial morphism $\beta : Sy^S \to [\mathbb{T}y, p]$ such that, for any section $\sigma : p \to y$, the morphism

$$Sy^S \xrightarrow{\beta} [\mathbb{T}y, p] \xrightarrow{[\mathbb{T}y,\sigma]} [\mathbb{T}y, y] \xrightarrow{\sim} y^{\mathbb{T}}$$

is a $\triangleleft$-comonoid homomorphism.

- β is equivalently a morphism $\mathbb{T}y \otimes Sy^S \to p$, and hence a pair (β^o, β^u), with
  - output map $\beta^o : \mathbb{T} \times S \to p(1)$,
  - update map $\beta^u : \sum_{t:\mathbb{T}} \sum_{s:S} p[\beta^o(t, s)] \to S$.
- Comonoid Sy^S is the codiscrete groupoid on S; $y^{\mathbb{T}}$ is the delooping $\mathbf{B}\mathbb{T}$ of $\mathbb{T}$;
- Comonoid homomorphism axiom enforces the monoid action condition:

  for any section $\sigma : p \to y$, the *closure* $\beta^\sigma(t) := S \xrightarrow{\beta^o(t)^*\sigma} \sum_{s:S} p[\beta^o(t, s)] \xrightarrow{\beta^u} S$

  constitutes a closed dynamical system on S

Examples of deterministic systems

Fix an interface polynomial p.

1. In discrete time, with $\mathbb{T} = \mathbb{N}$, a system β is again determined by its components at $t = 1$:
   - output $\beta^o : S \to p(1)$,
   - update $\beta^u : \sum_{s:S} p\big[\vartheta^o(s)\big] \to S$;
   - this is equivalently a p-coalgebra $\beta : S \to pS$ — so our systems are 'generalized p-coalgebras'.

2. Solutions to 'open ODEs' are systems with $\mathbb{T} = \mathbb{R}$:
   - suppose output $\vartheta^o : S \to p(1)$ and 'open' vector field $f : \sum_{s:S} p\big[\vartheta^o(s)\big] \to TS$, so that $\dfrac{\mathrm{d}s}{\mathrm{d}t} = f(s, x)$ for each $x : p\big[\vartheta^o(s)\big]$;
   - the solutions give us a family of maps $(s_0, x) \mapsto s_x(t)$;
   - and hence an update map $\vartheta^u : (t, s_0, x) \mapsto s_x(t)$;
   - such that pulling back along ϑ^o gives the closed system $s_0 \mapsto s_{\vartheta^o(s)}(t)$;
   - thus satisfying the action condition.

Deterministic open systems over p form a category

Recall that a morphism $f : \vartheta \to \psi$ of closed dynamical systems is a natural transformation:

$$\begin{array}{ccc} X & \xrightarrow{\vartheta(t)} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[\psi(t)]{} & Y \end{array}$$

i.e. a morphism $f : X \to Y$ of state spaces making the above squares commute for all $t : \mathbb{T}$.

Deterministic open systems over p form a category

We can extend the same idea to systems over p, using their closures.

A morphism $f : \vartheta \to \phi$ is a map $f : X \to Y$ of state spaces, such that these commute:

$$\begin{array}{ccccc} X & \xrightarrow{\vartheta^o(t)^*\sigma} & \sum_{x:X} p[\vartheta^o(t,x)] & \xrightarrow{\vartheta^u(t)} & X \\ {\scriptstyle f}\downarrow & & & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow[\psi^o(t)^*\sigma]{} & \sum_{y:Y} p[\psi^o(t,y)] & \xrightarrow[\psi^u(t)]{} & Y \end{array}$$

The identity on ϑ is just the identity map on its state space.

We denote the resulting category by $\mathbf{Coalg}^{\mathbb{T}}_{\mathcal{E}}(p)$.

$\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}$ is opindexed $\mathbf{Poly}_{\mathcal{E}} \to \mathbf{Cat}$

Each object of $\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}(p)$ is a morphism $Sy^S \to [\mathbb{T}y, p]$ for some S.

Given a morphism $\varphi : p \to q$, we can reindex covariantly:

$$Sy^S \to [\mathbb{T}y, p] \xrightarrow{[\mathbb{T}y,\varphi]} [\mathbb{T}y, q]$$

It is easy to show that, if $f : \beta \to \beta'$ is a morphism of systems, then the underlying map of state spaces gives a morphism $\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}(\varphi)(f) : \mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}(\varphi)(\beta) \to \mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}(\varphi)(\beta')$ after reindexing.

Hence we have an opindexed category $\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}} : \mathbf{Poly}_{\mathcal{E}} \to \mathbf{Cat}$.

Deterministic open systems are coalgebras

We have already seen that discrete-time systems are p-coalgebras. What about general time?

We can show that a morphism $Sy^S \to q$ is equivalently a map $S \to q(S)$: a q-coalgebra.

By setting $q = [\mathbb{T}y, p]$, we find: each system $Sy^S \to [\mathbb{T}y, p]$ is a coalgebra $S \to [\mathbb{T}y, p](S)$.

- but not all such coalgebras are 'systems' (they may not induce cofunctors)

Now what about stochastic / M-type systems?

- after all, closed M-systems are M-coalgebras ...

Stochastic and other systems

Given a monad $M : \mathcal{E} \to \mathcal{E}$, we can construct an analogous indexed category $\mathbf{Coalg}^{\mathbb{T}}_{M}$.

For each polynomial p, the objects are triples (S, β^o, β^u) with

- state space $S : \mathcal{E}$,
- output map $\beta^o : \mathbb{T} \times S \to p(1)$, and
- update map $\beta^u : \sum_{t:\mathbb{T}} \sum_{s:S} p[\beta^o(t,s)] \to MS$,

such that for each $\sigma : p \to y$, the closure

$$\beta^\sigma(t) : S \xrightarrow{\beta^o(t)^*\sigma} \sum_{s:S} p[\beta^o(t,s)] \xrightarrow{\beta^u} MS$$

induces a functor $\mathbf{B}\mathbb{T} \to \mathcal{K}\ell(M)$.

Stochastic and other systems

Given a monad $M : \mathcal{E} \to \mathcal{E}$, we can construct an analogous indexed category $\mathbf{Coalg}^{\mathbb{T}}_M$.

For each polynomial p, the objects are triples (S, β^o, β^u) with

- state space $S : \mathcal{E}$,
- output map $\beta^o : \mathbb{T} \times S \to p(1)$, and
- update map $\beta^u : \sum_{t:\mathbb{T}} \sum_{s:S} p[\beta^o(t,s)] \to MS$,

such that for each $\sigma : p \to y$, the closure

$$\beta^\sigma(t) : S \xrightarrow{\beta^o(t)^*\sigma} \sum_{s:S} p[\beta^o(t,s)] \xrightarrow{\beta^u} MS$$

induces a functor $\mathbf{B}\mathbb{T} \to \mathcal{K}\ell(M)$.

We can obtain $\mathbf{Coalg}^{\mathbb{T}}_M$ by instantiating def. (2.1) in a "cat. of M-polynomials", $\mathbf{Poly}_M$.

$\mathbf{Poly}_M$ has same objects as $\mathbf{Poly}_{\mathcal{E}}$; backwards maps are families of morphisms in $\mathcal{K}\ell(M)$.

$\mathbf{Poly}_M$ arises as a modified category of dependent lenses.

- It doesn't have $\prod$-types in general, or a universal internal hom.
- But it does have *some* $\prod$-types, and a 'deterministic' hom.

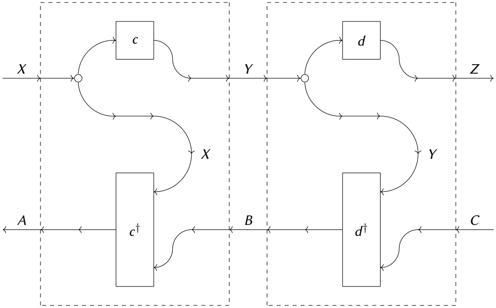
i.e., enough to make the definition work!

For details, see the paper :-)

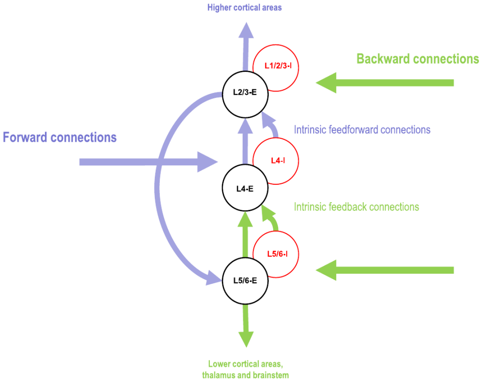
We can also make random dynamical systems work ... (but no time for that today!)

Monoidal bicategories of open dynamical systems

So, given $\mathbf{Coalg}^{\mathbb{T}}_{M}$, how to set up the semantics for the "predictive coding" functor?



$\mapsto$



Here, we need 'hierarchical' dynamical systems: 1-cells with type $(X, A) \to (Y, B)$ ...

Monoidal bicategories of open dynamical systems

The answer is to construct 'hom' polynomials, akin to $[-, =]$,
and then consider dynamical systems on these interfaces.

- This is a very similar approach to that of Shapiro and Spivak (which we'll hear about later), and effectively gives rise to a generalization of Spivak's operad **Org**. (See the papers for details!)

Monoidal bicategories of open dynamical systems

The answer is to construct 'hom' polynomials, akin to $[-, =]$, and then consider dynamical systems on these interfaces.

- This is a very similar approach to that of Shapiro and Spivak (which we'll hear about later), and effectively gives rise to a generalization of Spivak's operad **Org**. (See the papers for details!)

Example: for each pair of objects $A, B : \mathcal{C}$, define a polynomial $\{A, B\} := \mathcal{C}(A, B)\, y^{\mathcal{C}(1,A)}$. Then there is a natural composition morphism, $\mathrm{c}_{A,B,C} : \{A, B\} \otimes \{B, C\} \to \{A, C\}$.

- *N.b.*, when $\mathcal{C} = \mathcal{E}$, we have $\{A, B\} = [Ay, By]$.

We obtain a bicategory $\mathbf{Hier}|_{\mathcal{C}}$: the hom categories are given by $\mathbf{Coalg}^{\mathbb{T}}_{M}(\{A, B\})$; composition is given by $\mathrm{Coalg}^{\mathbb{T}}_{M}(\mathrm{c})$, and identities are the constant systems emitting identities.

Monoidal bicategories of open dynamical systems

The answer is to construct 'hom' polynomials, akin to $[-, =]$, and then consider dynamical systems on these interfaces.

- This is a very similar approach to that of Shapiro and Spivak (which we'll hear about later), and effectively gives rise to a generalization of Spivak's operad **Org**. (See the papers for details!)

Example: for each pair of objects $A, B : \mathcal{C}$, define a polynomial $\{A, B\} := \mathcal{C}(A, B)\, y^{\mathcal{C}(1,A)}$.
Then there is a natural composition morphism, $\mathsf{c}_{A,B,C} : \{A, B\} \otimes \{B, C\} \to \{A, C\}$.

- *N.b.*, when $\mathcal{C} = \mathcal{E}$, we have $\{A, B\} = [Ay, By]$.

We obtain a bicategory $\mathbf{Hier}|_{\mathcal{C}}$: the hom categories are given by $\mathbf{Coalg}^{\mathbb{T}}_{M}(\{A, B\})$; composition is given by $\mathrm{Coalg}^{\mathbb{T}}_{M}(\mathsf{c})$, and identities are the constant systems emitting identities.

And when $\mathcal{C}$ is a copy-discard category, then so is $\mathbf{Hier}|_{\mathcal{C}}$.

- This means we can give a 'dynamical Bayes rule' ...
- ... but we need to be careful to use *bisimulation* rather than strict equality.
    - (See the paper for details!)

Dynamical cybernetics and 'cilia'

Example: To get closer to **Org**, we can define hom-categories $\mathbf{Hier}(p, q) := \mathbf{Coalg}_M^{\mathbb{T}}([p, q])$.
This gives a monoidal bicategory **Hier**, which restricts to $\mathbf{Hier}|_{\mathcal{E}}$ on linear polynomials.
This is a bicategory of 'dynamical dependent lenses'.

Alternatively, consider other bidirectional or 'cybernetic' categories — such as Bayesian lenses ...

Dynamical cybernetics and 'cilia'

Example: the category of Bayesian lenses has as objects pairs (X, A) of objects in $\mathcal{E}$. Therefore define the polynomials $[\![Xy^A, Yy^B]\!] := \mathbf{BayesLens}\big((X,A),(Y,B)\big)\, y^{\mathcal{D}X\times B}$.

We obtain a similar composition morphism, and hence a bicategory **HierInf** with hom-categories $\mathbf{HierInf}(Xy^A, Yy^B) := \mathbf{Coalg}^{\mathbb{T}}_M\big([\![Xy^A, Yy^B]\!]\big)$.

A 1-cell $\vartheta : Xy^A \to Yy^B$ is then given by a tuple $(S, \vartheta^o_1, \vartheta^o_2, \vartheta^u)$:

- state space S;
- forwards output channel $\vartheta^o_1 : \mathbb{T} \times S \times X \to \mathcal{D}Y$;
- backwards output $\vartheta^o_2 : \mathbb{T} \times S \times \mathcal{D}X \times B \to \mathcal{D}A$;
- update $\vartheta^u : \mathbb{T} \times S \times \mathcal{D}X \times B \to \mathcal{D}S$.

This category is what we need for the dynamical semantics of predictive coding...

(... and will be in my forthcoming preprint *Compositional Active Inference II*)

Dynamical cybernetics and 'cilia'

Finally, we can do something similar for any category of optics.

Recall that for an optic $l : \Phi \to \Psi$, we have a notion of 'context':

$$\mathrm{Ctx}(l) := \int^{M:\mathbf{Optic}} \mathbf{Optic}(I, M \otimes \Phi) \times \mathbf{Optic}(M \otimes \Psi, I)$$

— (it's "everything needed to close off l".)

We define polynomials $\langle \Phi, \Psi \rangle := \sum_{l:\mathbf{Optic}(\Phi,\Psi)} y^{\mathrm{Ctx}(l)}$.

The yoga of optics gives us composition morphisms $\langle \Phi, \Psi \rangle \otimes \langle \Psi, \mathrm{X} \rangle \to \langle \Phi, \mathrm{X} \rangle$.
And hence a monoidal bicategory **Cilia** with hom-categories $\mathbf{Cilia}(\Phi, \Psi) := \mathbf{Coalg}_M^{\mathbb{T}}(\langle \Phi, \Psi \rangle)$.

- *N.b.*, 'cilia' are the structures that control the eye.

Such categories may be useful for building dynamical systems that *e.g.* play open games.

Conclusions

To recap:

- we constructed opindexed categories $\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}$ of deterministic dynamical systems in a (LCC) category $\mathcal{E}$, for a monoid $\mathbb{T}$ modelling time;
- we saw how to extend these to systems with different transition types, using a monad M;
- we saw how these systems generalize closed systems, and how they constitute coalgebras;
- and we constructed bicategories of dynamical systems for cybernetic applications.

But there are many open questions!

- What is the connection to the monoidal streams of Di Lavore *et al*?
- Is $\mathbf{Coalg}_{\mathcal{E}}^{\mathbb{T}}$ a topos? Can we translate ideas from coalgebraic logic?
- How does this coalgebraic framework relate to Myers' double-categorical framework?

Thanks for listening!

Bibliography

1. Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J.: Canonical microcircuits for predictive coding. Neuron 76(4), 695–711 (2012). DOI: 10.1016/j.neuron.2012.10.038

2. Ahman, D., and Uustalu, T.: Directed Containers as Categories. EPTCS 207, 2016, pp. 89-98 (2016). DOI: 10.4204/EPTCS.207.5

3. Spivak, D.I., and Niu, N.: Polynomial Functors: A General Theory of Interaction. (2021)