# Polynomial Functors and Shannon Entropy

David I. Spivak

TOPOS INSTITUTE

Symposium on Categorical Semantics of Entropy
2022 May 13

# Outline

# Why am I here?

We're here to learn from each other. But what is learning?

- Somehow out of all the information out there, some of it *sticks*.
- We develop frameworks by which to *store* information.
- I'm interested in how intelligence and learning function.
- So I study how knowledge is stored and transferred in databases and...
- ...how dynamical systems interact to adapt and learn (e.g. in DNNs).

# Why am I here?

We're here to learn from each other. But what is learning?

- Somehow out of all the information out there, some of it *sticks*.
- We develop frameworks by which to *store* information.
- I'm interested in how intelligence and learning function.
- So I study how knowledge is stored and transferred in databases and...
- ...how dynamical systems interact to adapt and learn (e.g. in DNNs).

Entropy has been put forward as an approach to intelligence and learning.

- Life can be understood as a *dissipative system*, spraying entropy.
- It does so while packing negentropy—organization—into itself.
- Polani's *empowerment* and Freer's *causal entropic forces*...
- ...are entropy-based approaches to intelligent behavior.

# Why am I here?

We're here to learn from each other. But what is learning?

- Somehow out of all the information out there, some of it *sticks*.
- We develop frameworks by which to *store* information.
- I'm interested in how intelligence and learning function.
- So I study how knowledge is stored and transferred in databases and...
- ...how dynamical systems interact to adapt and learn (e.g. in DNNs).

Entropy has been put forward as an approach to intelligence and learning.

- Life can be understood as a *dissipative system*, spraying entropy.
- It does so while packing negentropy—organization—into itself.
- Polani's *empowerment* and Freer's *causal entropic forces*...
- ...are entropy-based approaches to intelligent behavior.

I only seem to understand things when they're written categorically.

- I've been trying to understand "what entropy really is."
- The Baez-Fritz-Leinster conception of entropy is great,...
- ...but I want to connect it in with dynamical systems or databases.

# The overwhelming abundance of Poly

In January 2020 I fell in love with a category called **Poly**.

- Its applications subsume everything I'd done with categ'l databases...
- ...and everything I'd done with interacting dynamical systems.
- It's used in functional programming, type theory, higher cat'y theory.

# The overwhelming abundance of Poly

In January 2020 I fell in love with a category called **Poly**.

- Its applications subsume everything I'd done with categ'l databases...
- ...and everything I'd done with interacting dynamical systems.
- It's used in functional programming, type theory, higher cat'y theory.

But it's not just very applicable, it's also very highly-structured.

- Coproducts and products that agree with usual polynomial arithmetic;
- All limits and colimits;
- At least three orthogonal factorization systems;
- A symmetric monoidal structure $\otimes$ distributing over $+$;
- A cartesian closure $q^p$ and monoidal closure $[p, q]$ for $\otimes$;
- Another nonsymmetric monoidal structure $\triangleleft$ that's duoidal with $\otimes$;
- A left (Meyers?) $\triangleleft$-coclosure $\left[\begin{smallmatrix}-\\-\end{smallmatrix}\right]$, meaning $\mathbf{Poly}(p, q \triangleleft r) \cong \mathbf{Poly}(\left[\begin{smallmatrix}r\\p\end{smallmatrix}\right], q)$;
- An indexed right $\triangleleft$-coclosure, i.e. $\mathbf{Poly}(p, q \triangleleft r) \cong \sum_{f\colon\, p(1)\to q(1)} \mathbf{Poly}(p \overset{f}{\frown} q, r)$;
- An indexed right $\otimes$-coclosure (Niu?), i.e. $\mathbf{Poly}(p, q \otimes r) \cong \sum_{f\colon\, p(1)\to q(1)} \mathbf{Poly}(p \overset{f}{\nearrow} q, r)$;
- At least eight monoidal structures in total;
- $\triangleleft$-monoids generalize $\Sigma$-free operads;
- $\triangleleft$-comonoids are exactly categories; bicomodules are data migrations.

See "A reference for categorical structures on **Poly**", arXiv: 2202.00534

# Entropy in terms of Poly

I now use the **Poly**-ecosystem to structure my thinking.

- The abundance of structure lets me track my mental moves.
- I can check the resulting formulation using concrete examples.
- So now I try to do *everything* in **Poly**, e.g. thinking about entropy.

# Entropy in terms of Poly

I now use the **Poly**-ecosystem to structure my thinking.

- The abundance of structure lets me track my mental moves.
- I can check the resulting formulation using concrete examples.
- So now I try to do *everything* in **Poly**, e.g. thinking about entropy.

So today I'll tell you how entropy looks from the **Poly** point of view.

- I'll show how to think of objects in **Poly** as empirical distributions.
- I'll show that there are distributive monoidal functors

$$\mathbf{Poly}^{\mathbf{Cart}} \xrightarrow{p \mapsto \dot{p}y} \mathbf{Poly} \xrightarrow{p \mapsto (p(1), \Gamma(p))} \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$$

  sending $p \in \mathbf{Poly}^{\mathbf{Cart}}$ to an invariant $\hbar(p) := (A, B) \in \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$.
- The Shannon entropy can then be extracted: $H(p) = \log(A^A/B)/A$.
- Properties of entropy follow from the distributive monoidality of $\hbar$.

# Plan

The plan for the rest of the time is as follows:

- Give background on polynomial functors.
- Explain $\hbar\colon \mathbf{Poly}^{\mathbf{Cart}} \to \mathbf{Set} \times \mathbf{Set}^{\mathsf{op}}$ and its relation to entropy.
- Talk about generalizations and future work.
- Conclude.

# Outline

# Poly for experts

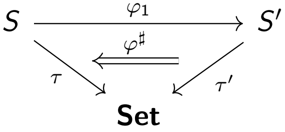
What I'll call the category **Poly** has many names.

- The free completely distributive category on one object;
- The free coproduct completion of $\mathbf{Set}^{\mathsf{op}}$;
- The full subcategory of $[\mathbf{Set}, \mathbf{Set}]$ spanned by...
  ...functors that preserve connected limits;
- The full subcategory of $[\mathbf{Set}, \mathbf{Set}]$ spanned by coproducts of repr'bles;

# Poly for experts

What I'll call the category **Poly** has many names.

- The free completely distributive category on one object;
- The free coproduct completion of $\mathbf{Set}^{\mathrm{op}}$;
- The full subcategory of $[\mathbf{Set}, \mathbf{Set}]$ spanned by...
  ...functors that preserve connected limits;
- The full subcategory of $[\mathbf{Set}, \mathbf{Set}]$ spanned by coproducts of repr'bles;
- The category of *typed sets* and colax maps between them.
  - Objects: pairs $(S, \tau)$, where $S \in \mathbf{Set}$ and $\tau \colon S \to \mathbf{Set}$.
  - Morphisms $(S, \tau) \xrightarrow{\varphi} (S', \tau')$: pairs $(\varphi_1, \varphi^\sharp)$, where



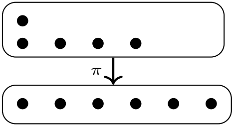

But let's make this easier.

# What is a polynomial?

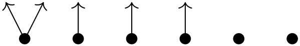
**Algebraic**

$$y^2 + 3y + 2$$

**Bundle**




**Corolla forest**

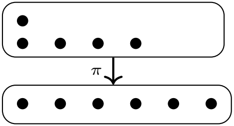


# What is a polynomial?

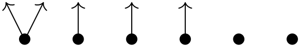
**Algebraic**

$y^2 + 3y + 2$

**Bundle**




**Corolla forest**



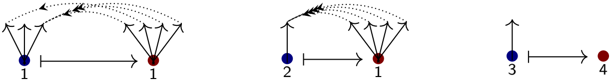
You can think of the bundle as a empirical distribution:

- The first outcome was drawn twice; the next three once; the rest never.
- It corresponds to the distribution $(\frac{2}{5}, \quad \frac{1}{5}, \quad \frac{1}{5}, \quad \frac{1}{5}, \quad 0, \quad 0)$.

# What is a morphism of polynomials?

Let $p \coloneqq y^3 + 2y$ and $q \coloneqq y^4 + y^2 + 2$


A morphism $p \xrightarrow{\varphi} q$ sends $p$-outcomes to $q$-outcomes, interpreting draws:



# The category of polynomials

Easiest description: **Poly** = "sums of representables functors **Set** → **Set**".

- For any set $S$, let $y^S := \mathbf{Set}(S, -)$, the functor *represented* by $S$.
- Def: a polynomial is a sum $p = \sum_{i \in I} y^{p[i]}$ of representable functors.
- Def: a morphism of polynomials is a natural transformation.
- In **Poly**, usual $+$ is the coproduct and usual $\times$ is the product.

# The category of polynomials

Easiest description: **Poly** = "sums of representables functors **Set** → **Set**".

- For any set $S$, let $y^S := \mathbf{Set}(S, -)$, the functor *represented* by $S$.
- Def: a polynomial is a sum $p = \sum_{i \in I} y^{p[i]}$ of representable functors.
- Def: a morphism of polynomials is a natural transformation.
- In **Poly**, usual $+$ is the coproduct and usual $\times$ is the product.

We will need a wide subcategory $\mathbf{Poly}^{\mathbf{Cart}} \subseteq \mathbf{Poly}$.

- Same objects, but morphisms $p \xrightarrow{\varphi} q$ are *cartesian natural transform's*;
- ...i.e. for any function $S \to T$, the naturality square is a pullback.
- Equivalently, for each outcome $i \in p(1)$ the interpretation map

$$q[\varphi(i)] \cong p[i]$$

is a bijection. Example: there are 24 cartesian maps $y^4 \to y^4 + y^3$.

# Notation

We said that a polynomial is a sum of representable functors

$$p \cong \sum_{i \in I} y^{p[i]}.$$

But note that $I \cong p(1)$. So we can write

$$p \cong \sum_{i \in p(1)} y^{p[i]}.$$

So $p(1)$ is the *outcome-set*, and elements of $p[i]$ are *draws* of outcome $i$.

# Fundamental invariants

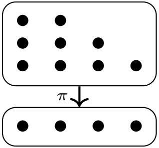
We will be interested in two fundamental invariants of a polynomial.

- From the bundle POV, these would be *base* and *global sections*.
- So if $p$ is represented by $E \to B$, these are $B$ and $\mathbf{Set}_{/B}(B, E)$.
- In terms of polynomials these are

$$p(1) \cong \mathbf{Poly}(y, p) \qquad \text{and} \qquad \Gamma(p) := \mathbf{Poly}(p, y).$$

- E.g. for the following bundle these are $p(1) \cong 4$ and $\Gamma(p) \cong 18$.

$$2y^3 + y^2 + y \quad =$$




has 4 base elements
and 18 global sections

# Fundamental invariants

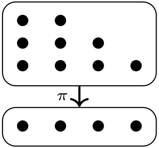
We will be interested in two fundamental invariants of a polynomial.

- From the bundle POV, these would be *base* and *global sections*.
- So if $p$ is represented by $E \to B$, these are $B$ and $\mathbf{Set}_{/B}(B, E)$.
- In terms of polynomials these are

$$p(1) \cong \mathbf{Poly}(y, p) \qquad \text{and} \qquad \Gamma(p) := \mathbf{Poly}(p, y).$$

- E.g. for the following bundle these are $p(1) \cong 4$ and $\Gamma(p) \cong 18$.

$$2y^3 + y^2 + y \quad =$$




has 4 base elements
and 18 global sections

These are functorial in opposite directions:

$$\mathbf{Poly} \xrightarrow{p \mapsto (p(1), \Gamma(p))} \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$$

In fact, this functor is a left adjoint, but we won't need that.

# The distributive monoidal structure $(\mathbf{Poly}, 0, +, y, \otimes)$

The category **Poly** is distributive monoidal.

- The usual sum of two polynomials is their coproduct; 0 is initial.
- The usual product is the cartesian product too, but we won't use this.
- There is another operation $\otimes$ called *Dirichlet product*. Formula:

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

These are very simple bundle-wise: sum & product of base and total space:

$$\begin{array}{c} E_1 \\ \downarrow \\ B_1 \end{array} + \begin{array}{c} E_2 \\ \downarrow \\ B_2 \end{array} = \begin{array}{c} E_1 + E_2 \\ \downarrow \\ B_1 + B_2 \end{array} \quad \text{and} \quad \begin{array}{c} E_1 \\ \downarrow \\ B_1 \end{array} \otimes \begin{array}{c} E_2 \\ \downarrow \\ B_2 \end{array} = \begin{array}{c} E_1 \times E_2 \\ \downarrow \\ B_1 \times B_2 \end{array}$$

# The distributive monoidal structure $(\mathbf{Poly}, 0, +, y, \otimes)$

The category **Poly** is distributive monoidal.

- The usual sum of two polynomials is their coproduct; 0 is initial.
- The usual product is the cartesian product too, but we won't use this.
- There is another operation $\otimes$ called *Dirichlet product*. Formula:

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

These are very simple bundle-wise: sum & product of base and total space:

$$\begin{array}{c} E_1 \\ \downarrow \\ B_1 \end{array} + \begin{array}{c} E_2 \\ \downarrow \\ B_2 \end{array} = \begin{array}{c} E_1 + E_2 \\ \downarrow \\ B_1 + B_2 \end{array} \qquad \text{and} \qquad \begin{array}{c} E_1 \\ \downarrow \\ B_1 \end{array} \otimes \begin{array}{c} E_2 \\ \downarrow \\ B_2 \end{array} = \begin{array}{c} E_1 \times E_2 \\ \downarrow \\ B_1 \times B_2 \end{array}$$

- These clearly distribute: $p \otimes (q_1 + q_2) \cong (p \otimes q_1) + (p \otimes q_2)$.
- Soon we'll see how the fundamental invariants respect these oper'ns.

# Derivatives and the total space

The derivative of a polynomial functor is another polynomial functor.

- Write $\dot{p}$ for the derivative with respect to $y$.
- In fact, we will be much more interested in $\dot{p}y$.

$$\dot{p} = \sum_{i \in p(1)} \sum_{d \in p[i]} y^{p[i] \setminus \{d\}} \qquad \text{and} \qquad \dot{p}y \cong \sum_{i \in p(1)} p[i] y^{p[i]}$$

# Derivatives and the total space

The derivative of a polynomial functor is another polynomial functor.

- Write $\dot{p}$ for the derivative with respect to $y$.
- In fact, we will be much more interested in $\dot{p}y$.

$$\dot{p} = \sum_{i \in p(1)} \sum_{d \in p[i]} y^{p[i] \setminus \{d\}} \qquad \text{and} \qquad \dot{p}y \cong \sum_{i \in p(1)} p[i] y^{p[i]}$$

Neither of these is functorial in **Poly**, but *they are functorial* in $\mathbf{Poly}^{\mathbf{Cart}}$.

$$\mathbf{Poly}^{\mathbf{Cart}} \xrightarrow{p \mapsto \dot{p}y} \mathbf{Poly}^{\mathbf{Cart}}$$

# Derivatives and the total space

The derivative of a polynomial functor is another polynomial functor.

- Write $\dot{p}$ for the derivative with respect to $y$.
- In fact, we will be much more interested in $\dot{p}y$.

$$\dot{p} = \sum_{i \in p(1)} \sum_{d \in p[i]} y^{p[i] \setminus \{d\}} \qquad \text{and} \qquad \dot{p}y \cong \sum_{i \in p(1)} p[i] y^{p[i]}$$

Neither of these is functorial in **Poly**, but *they are functorial* in $\mathbf{Poly}^{\mathbf{Cart}}$.

$$\mathbf{Poly}^{\mathbf{Cart}} \xrightarrow{p \mapsto \dot{p}y} \mathbf{Poly}^{\mathbf{Cart}}$$

In fact $p \mapsto \dot{p}y$ is a comonad on $\mathbf{Poly}^{\mathbf{Cart}}$.

- I see the counit $\dot{p}y \to p$ as “how **Poly** thinks of $p$ as a bundle.”
- In that way $p$ is the *base*. So let’s call $p \mapsto \dot{p}y$ the *total space functor*.

# Bifibration Poly $\to$ Set

The last theory we'll need is the bifibration **Poly** $\to$ **Set**.

- The functor $p \mapsto p(1)$ is both a fibration and an op-fibration.
- In fact it's even more: a distributive monoidal $*$-bifibration!

Down to earth what does this mean? Let $p$ be a polynomial and $A$ a set.

- For any function $f\colon A \to p(1)$ we can take the pullback

$$\begin{array}{ccc} f^*p & \longrightarrow & p \\ \downarrow & \lrcorner & \downarrow \\ A & \xrightarrow[f]{} & p(1) \end{array}$$

# Bifibration Poly $\to$ Set

The last theory we'll need is the bifibration **Poly** $\to$ **Set**.

- The functor $p \mapsto p(1)$ is both a fibration and an op-fibration.
- In fact it's even more: a distributive monoidal $*$-bifibration!

Down to earth what does this mean? Let $p$ be a polynomial and $A$ a set.

- For any function $f\colon A \to p(1)$ we can take the pullback

$$\begin{array}{ccc} f^*p & \longrightarrow & p \\ \downarrow & \lrcorner & \downarrow \\ A & \xrightarrow[f]{} & p(1) \end{array}$$

- So $f^*$ takes polys with outcome-set $p(1)$ to those with outcome-set $A$.
- This operation has both a left adjoint $f_!$ and a right adjoint $f_*$.

$$f_!p := \sum_{b \in B} y^{\prod_{a \mapsto b} p[a]} \qquad \text{and} \qquad f_*p := \sum_{b \in B} y^{\sum_{a \mapsto b} p[a]}$$

- I.e., for any $f\colon A \to B$, we have $\mathbf{Poly}_A(f^*q, p) \cong \mathbf{Poly}_B(q, f_*p)$.

As we lump outcomes together, we add up the draws.

# Outline

# Total space as distributive

We're now ready to get to work on how all this relates to entropy.

- The approach is to extract two invariant sets from any polynomial.
- This process is "good" in that it is a distributive monoidal functor.
- We'll extract the extensive and intensive entropies from these.

# Total space as distributive

We're now ready to get to work on how all this relates to entropy.

- The approach is to extract two invariant sets from any polynomial.
- This process is "good" in that it is a distributive monoidal functor.
- We'll extract the extensive and intensive entropies from these.

So first, we want to see that $p \mapsto \dot{p}y$ is distributive monoidal.

- The derivative is linear, $(p \dot{+} q) = \dot{p} + \dot{q}$, and so is $p \mapsto py$.
- So $p \mapsto \dot{p}y$ preserves coproducts. What about $\otimes$?

$$
\begin{aligned}
(\dot{p}y) \otimes (\dot{q}y) &\cong \sum_{i \in p(1)} p[i] y^{p[i]} \otimes \sum_{j \in q(1)} q[j] y^{q[j]} \\
&\cong \sum_{(i,j) \in p(1) \times q(1)} p[i] \times q[j] y^{p[i] \times q[j]} \\
&\cong (p \dot{\otimes} q) y
\end{aligned}
$$

# Total space as distributive

We're now ready to get to work on how all this relates to entropy.

- The approach is to extract two invariant sets from any polynomial.
- This process is "good" in that it is a distributive monoidal functor.
- We'll extract the extensive and intensive entropies from these.

So first, we want to see that $p \mapsto \dot{p}y$ is distributive monoidal.

- The derivative is linear, $(\dot{p + q}) = \dot{p} + \dot{q}$, and so is $p \mapsto py$.
- So $p \mapsto \dot{p}y$ preserves coproducts. What about $\otimes$?

$$\begin{aligned}(\dot{p}y) \otimes (\dot{q}y) &\cong \sum_{i \in p(1)} p[i] y^{p[i]} \otimes \sum_{j \in q(1)} q[j] y^{q[j]} \\ &\cong \sum_{(i,j) \in p(1) \times q(1)} p[i] \times q[j] y^{p[i] \times q[j]} \\ &\cong (\dot{p \otimes q})y\end{aligned}$$

So $(\mathbf{Poly}^{\mathbf{Cart}}, 0, +, y, \otimes) \xrightarrow{p \mapsto \dot{p}y} (\mathbf{Poly}, 0, +, y, \otimes)$ is distributive monoidal.

# Fundamental invariants as distributive

The fundamental invariants $p \mapsto (p(1), \Gamma(p))$ are also distributive

$$(\mathbf{Poly}, 0, +, y, \otimes) \xrightarrow{p \mapsto (p(1), \Gamma(p))} (\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}, (0, 1), +, (1, 1), \otimes)$$

But what exactly is all this structure on $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$?

# Fundamental invariants as distributive

The fundamental invariants $p \mapsto (p(1), \Gamma(p))$ are also distributive

$$(\mathbf{Poly}, 0, +, y, \otimes) \xrightarrow{p \mapsto (p(1), \Gamma(p))} (\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}, (0,1), +, (1,1), \otimes)$$

But what exactly is all this structure on $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$?

- $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ has coproducts: $(A_1, B_1) + (A_2, B_2) \cong (A_1 + A_2, B_1 \times B_2)$.
- It has another symmetric monoidal structure with unit $(1, 1)$:

$$(A_1, B_1) \otimes (A_2, B_2) := \left(A_1 \times A_2, B_1^{A_2} \times B_2^{A_1}\right)$$

- And these distribute "because" $B^{A_1 + A_2}(B_1 B_2)^A \cong (B^{A_1} B_1^A)(B^{A_2} B_2^A)$.

# Fundamental invariants as distributive

The fundamental invariants $p \mapsto (p(1), \Gamma(p))$ are also distributive

$$(\mathbf{Poly}, 0, +, y, \otimes) \xrightarrow{p \mapsto (p(1), \Gamma(p))} (\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}, (0,1), +, (1,1), \otimes)$$

But what exactly is all this structure on $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$?

- $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ has coproducts: $(A_1, B_1)+(A_2, B_2) \cong (A_1+A_2, B_1 \times B_2)$.
- It has another symmetric monoidal structure with unit $(1,1)$:

$$(A_1, B_1) \otimes (A_2, B_2) := \left(A_1 \times A_2, B_1^{A_2} \times B_2^{A_1}\right)$$

- And these distribute "because" $B^{A_1+A_2}(B_1 B_2)^A \cong (B^{A_1} B_1^A)(B^{A_2} B_2^A)$.

Why do the fundamental invariants (as a pair) preserve $+$ and $\otimes$?

- We have $(p+q)(1) \cong p(1) + q(1)$ and $\Gamma(p+q) \cong \Gamma(p) \times \Gamma(q)$.
- This says they preserve $+$. One also checks they preserve $\otimes$:

$$(p \otimes q)(1) \cong p(1) \times q(1) \qquad \text{and} \qquad \Gamma(p \otimes q) \cong \Gamma(p)^{q(1)} \times \Gamma(q)^{p(1)}$$

# Taking stock

Let's denote the composite of our distributive functors by $\hbar$:

$$\hbar : \mathbf{Poly}^{\mathbf{Cart}} \xrightarrow{p \mapsto \dot{p}y} \mathbf{Poly} \xrightarrow{p \mapsto (p(1), \Gamma(p))} \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$$

- The claim is that $\hbar$ extracts everything you need to calculate entropy.
- Preserving $+$ and $\otimes$ gives us properties of entropy.

# Taking stock

Let's denote the composite of our distributive functors by $\hbar$:

$$\mathbf{Poly}^{\mathsf{Cart}} \xrightarrow{p \mapsto \dot{p}y} \mathbf{Poly} \xrightarrow{p \mapsto (p(1), \Gamma(p))} \mathbf{Set} \times \mathbf{Set}^{\mathsf{op}}$$

$$\mathbf{Poly}^{\mathsf{Cart}} \xrightarrow{\hbar} \mathbf{Set} \times \mathbf{Set}^{\mathsf{op}}$$

- The claim is that $\hbar$ extracts everything you need to calculate entropy.
- Preserving $+$ and $\otimes$ gives us properties of entropy.

Define a real number $L(A, B) := \frac{\log(\#A^{\#A}) - \log(\#B)}{\#A}$. Then:

**Theorem**

*Let $p$ be a polynomial, considered as a probability distribution $P$, and let $H(P)$ be its Shannon entropy. Then we have*

$$H(P) = L(\hbar(p))$$

# The categorical partition function and entropy

I'm unfamiliar with the thermo picture. Joint with James Dama:

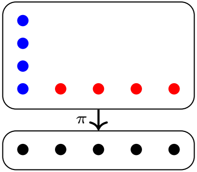
- One should think of Shannon entropy as an *entropy density*.
- The thermo picture defines a *partition function* $\Omega$ for distributions.
- For $p \in \textbf{Poly}$ with $\hbar(p) = (A, B)$, this would be $\Omega_p := \frac{A^A}{B}$.
- Then the extensive entropy of $p$ is given by $\mathrm{E}(p) := \log \Omega(p)$.
- And the Shannon entropy of $p$ is the density $H(p) := \mathrm{E}(p)/A$.

# The categorical partition function and entropy

I'm unfamiliar with the thermo picture. Joint with James Dama:

- One should think of Shannon entropy as an *entropy density*.
- The thermo picture defines a *partition function* $\Omega$ for distributions.
- For $p \in \mathbf{Poly}$ with $\hbar(p) = (A, B)$, this would be $\Omega_p := \frac{A^A}{B}$.
- Then the extensive entropy of $p$ is given by $\mathrm{E}(p) := \log \Omega(p)$.
- And the Shannon entropy of $p$ is the density $H(p) := \mathrm{E}(p)/A$.

For example consider the bundle for $p := y^4 + 4y^1$:




- We find $\dot{p}y = 4y^4 + 4y$, so $\hbar(p) = (4 + 4, 4^4) = (8, 4^4)$.
- So $\Omega_p = \frac{8^8}{4^4}$, Ext've: $\mathrm{E}(p) = \log \Omega_p = 16$, Shannon: $H(p) = 16/8 = 2$.

# Consequences of distributivity

The fact that $h$ preserves $\otimes$ and properties of log immediately give us

$$H(p \otimes q) = H(p) + H(q).$$

# Consequences of distributivity

The fact that $\hbar$ preserves $\otimes$ and properties of log immediately give us

$$H(p \otimes q) = H(p) + H(q).$$

The fact that $\hbar$ preserves $+$ is more subtle. We see it better in $\Omega$ and E.

- Again joint with James Dama.
- Suppose you write $p$ as a sum, $p := \sum_{a \in A} p_a$.
- This is the same as giving a function $f \colon p(1) \to A$.
- Recall that from this we get $f_* p \in \mathbf{Poly}_A$, lumping distributions.

# Consequences of distributivity

The fact that $\hbar$ preserves $\otimes$ and properties of log immediately give us

$$H(p \otimes q) = H(p) + H(q).$$

The fact that $\hbar$ preserves $+$ is more subtle. We see it better in $\Omega$ and E.

- Again joint with James Dama.
- Suppose you write $p$ as a sum, $p := \sum_{a \in A} p_a$.
- This is the same as giving a function $f \colon p(1) \to A$.
- Recall that from this we get $f_* p \in \mathbf{Poly}_A$, lumping distributions.

It follows from the fact that $\hbar$ preserves sums that:

$$\Omega_p = \Omega_{f_* p} \times \prod_{a \in A} \Omega_{p_a} \qquad \text{and} \qquad \mathrm{E}(p) = \mathrm{E}(f_* p) + \sum_{a \in A} \mathrm{E}(p_a)$$

The usual "chain rule" for Shannon entropy follows directly from this.

# A geometric viewpoint on Shannon entropy

Think of objects $(A, B) \in \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ as representing *formal rectangles*.

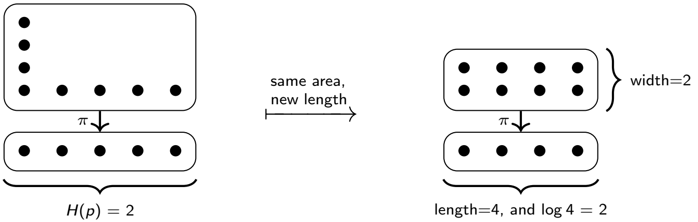
- Here $A$ is its area, $\sqrt[A]{B}$ is its width, and $A/\sqrt[A]{B}$ is its length.
- Adding two rectangles $(A_1, B_1) + (A_2, B_2) = (A_1 + A_2, B_1 \times B_2)$...
- ...add the areas and take the weighted geometric mean of the widths.
- Multiplying two rectangles $(A_1, B_1) \otimes (A_2, B_2) = (A_1 A_2, B_1^{A_2} B_2^{A_2})$...
- ...you multiply the areas and multiply the widths.
- The log-length $\log(A/\sqrt[A]{B})$ of the rectangle is the Shannon entropy.

# A geometric viewpoint on Shannon entropy

Think of objects $(A, B) \in \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ as representing *formal rectangles*.

- Here $A$ is its area, $\sqrt[A]{B}$ is its width, and $A/\sqrt[A]{B}$ is its length.
- Adding two rectangles $(A_1, B_1) + (A_2, B_2) = (A_1 + A_2, B_1 \times B_2)$...
- ...add the areas and take the weighted geometric mean of the widths.
- Multiplying two rectangles $(A_1, B_1) \otimes (A_2, B_2) = (A_1 A_2, B_1^{A_2} B_2^{A_2})$...
- ...you multiply the areas and multiply the widths.
- The log-length $\log(A/\sqrt[A]{B})$ of the rectangle is the Shannon entropy.

Let $p := y^4 + 4y$, so $\hbar(p) = (8, 4^4)$. Width$=\sqrt[8]{4^4} = 2$, length$=8/2 = 4$.

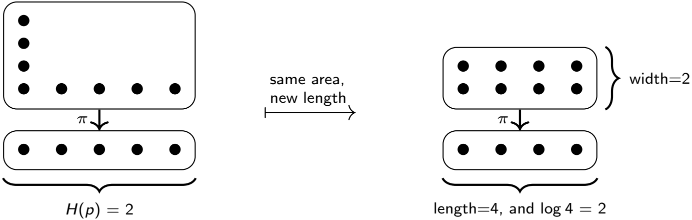


# A geometric viewpoint on Shannon entropy

Think of objects $(A, B) \in \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ as representing *formal rectangles*.

- Here $A$ is its area, $\sqrt[A]{B}$ is its width, and $A/\sqrt[A]{B}$ is its length.
- Adding two rectangles $(A_1, B_1) + (A_2, B_2) = (A_1 + A_2, B_1 \times B_2)$...
- ...add the areas and take the weighted geometric mean of the widths.
- Multiplying two rectangles $(A_1, B_1) \otimes (A_2, B_2) = (A_1 A_2, B_1^{A_2} B_2^{A_2})$...
- ...you multiply the areas and multiply the widths.
- The log-length $\log(A/\sqrt[A]{B})$ of the rectangle is the Shannon entropy.

Let $p := y^4 + 4y$, so $\hbar(p) = (8, 4^4)$. Width$=\sqrt[8]{4^4} = 2$, length$=8/2 = 4$.




If $q = Ly^W$ was rectangular to begin with, it'll stay that way.

# Outline

# Functoriality?

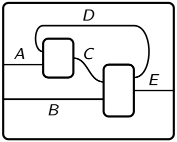
One of the big questions for me is: what does functoriality buy you?

- The distributivity of $\hbar\colon \mathbf{Poly}^{\mathbf{Cart}} \to \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ means something.
- It gives us well-known facts about entropy and entropy density.
- But what about the fact that $\hbar$ is functorial?
- Logarithms have no clue about what maps in $\mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$ mean.

# Entropy and dynamics?

Returning to my goals, I'd like to understand learning.

- If entropy will be involved, I want it to be about dynamical systems.
- The groupoid $\dot{p}y$ is kind of dynamic: little $p[i]$'s spinning around.
- But what about the *point* of Shannon entropy: communication?




- Shouldn't we be able to see Huffman coding or something here...?
- What about "empowerment" or "causal entropic forces"?

# Outline

# Summary

Entropy and **Poly** are amazing and have overlapping applications.

- Entropy is well-known throughout the scientific and technical world.
- **Poly** is the most highly-structured category I've ever seen.
- Both have applications to interacting dynamical systems.

# Summary

Entropy and **Poly** are amazing and have overlapping applications.

- Entropy is well-known throughout the scientific and technical world.
- **Poly** is the most highly-structured category I've ever seen.
- Both have applications to interacting dynamical systems.

There is (at least one) interpretation of entropy within **Poly**.

- Objects in **Poly** can be viewed as empirical distributions.
- There is a $(+, \otimes)$-preserving functor $\hbar \colon \mathbf{Poly}^{\mathbf{Cart}} \to \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$.
- If $\hbar(p) = (A, B)$ then $H(p) = \log(A/\sqrt[A]{B})/A$.
- So all the entropy-relevant data of $p$ is encapsulated in two sets.

# Summary

Entropy and **Poly** are amazing and have overlapping applications.

- Entropy is well-known throughout the scientific and technical world.
- **Poly** is the most highly-structured category I've ever seen.
- Both have applications to interacting dynamical systems.

There is (at least one) interpretation of entropy within **Poly**.

- Objects in **Poly** can be viewed as empirical distributions.
- There is a $(+, \otimes)$-preserving functor $\hbar\colon \mathbf{Poly}^{\mathbf{Cart}} \to \mathbf{Set} \times \mathbf{Set}^{\mathrm{op}}$.
- If $\hbar(p) = (A, B)$ then $H(p) = \log(A/\sqrt[A]{B})/A$.
- So all the entropy-relevant data of $p$ is encapsulated in two sets.

Entropy still feels somehow foreign to me.

- Hopefully, having different categorifications will help clarify it.
- I still have hope that it will bond with the dynamics of **Poly**.

*Thanks! Comments and questions welcome...*