Dynamic PROPs for Networked Learners

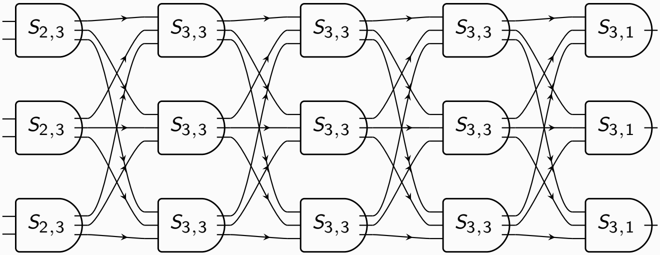
Brandon T. Shapiro* and David I. Spivak

CALCO 2023

TOPOS INSTITUTE

Outline

1. Nested dynamic structures
2. Polynomials and wiring diagrams
3. Dynamics: polynomial coalgebras
4. Nesting and networks: operad and PROP structure
5. A dynamic PROP for deep learning
6. Vistas

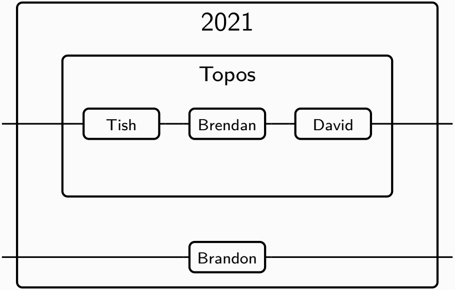


Nested dynamic structures

- How I joined Topos Institute

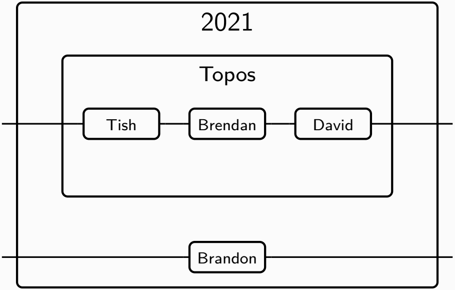
Nested dynamic structures

- How I joined Topos Institute



Nested dynamic structures

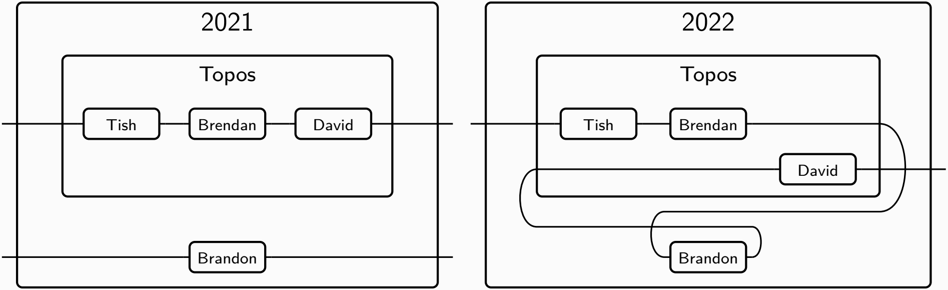
- How I joined Topos Institute



(Not an accurate representation of Topos Institute's internal structure)

Nested dynamic structures

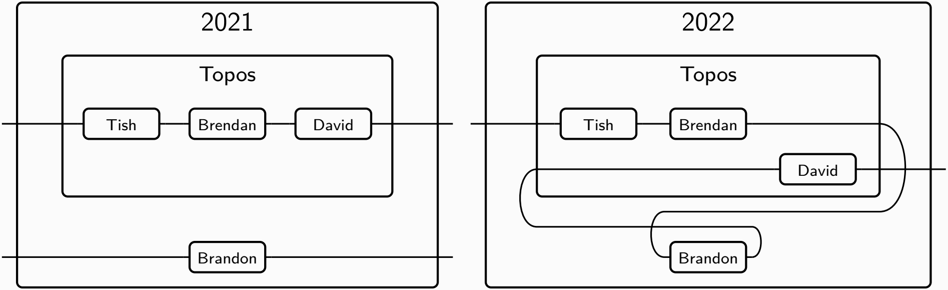
- How I joined Topos Institute

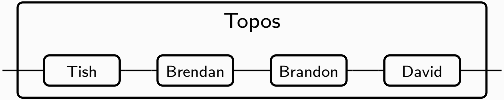


(Not an accurate representation of Topos Institute's internal structure)

Nested dynamic structures

- How I joined Topos Institute





(Not an accurate representation of Topos Institute's internal structure)

Polynomials and wiring diagrams

Polynomials and wiring diagrams

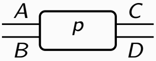
- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$

Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$



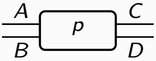
Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$

$$q = EFy^{C}$$

$$r = FGy^{A}$$



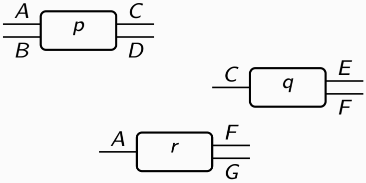
Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$

$$q = EFy^{C}$$

$$r = FGy^{A}$$



Polynomials and wiring diagrams

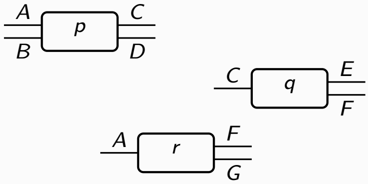
- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$

$$q = EFy^{C}$$

$$r = FGy^{A}$$

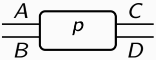
$$p \otimes q = CDEFy^{ABC}$$



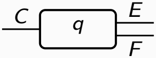
Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

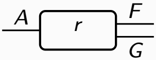
$$p = CDy^{AB}$$



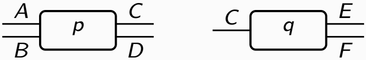
$$q = EFy^{C}$$



$$r = FGy^{A}$$



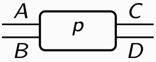
$$p \otimes q = CDEFy^{ABC}$$



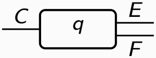
Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

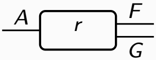
$$p = CDy^{AB}$$



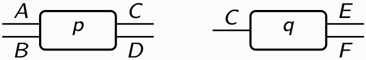
$$q = EFy^{C}$$



$$r = FGy^{A}$$



$$p \otimes q = CDEFy^{ABC}$$



- Polynomials form a category where a morphism $p \otimes q \to r$ consists of functions

Polynomials and wiring diagrams

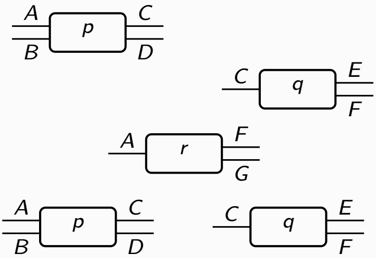
- Let A, B, C, D, E, F, G be sets, and consider the polynomials

$$p = CDy^{AB}$$

$$q = EFy^{C}$$

$$r = FGy^{A}$$

$$p \otimes q = CDEFy^{ABC}$$



- Polynomials form a category where a morphism $p \otimes q \to r$ consists of functions

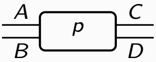
$$CDEF \to FG$$

$$(CDEF)A \to ABC$$

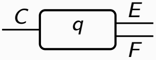
Polynomials and wiring diagrams

- Let A, B, C, D, E, F, G be sets, and consider the polynomials

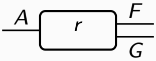
$$p = CDy^{AB}$$



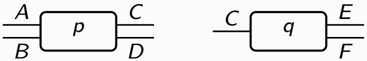
$$q = EFy^{C}$$



$$r = FGy^{A}$$



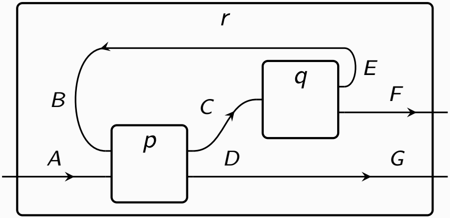
$$p \otimes q = CDEFy^{ABC}$$



- Polynomials form a category where a morphism $p \otimes q \to r$ consists of functions

$$CDEF \to FG$$

$$(CDEF)A \to ABC$$



Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- For $p(S) = A \times S^B$, each state is assigned an element of A and a function $B \to S$ which updates the state

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of “states” with a function $S \to p(S)$
- For $p(S) = A \times S^B$, each state is assigned an element of A and a function $B \to S$ which updates the state
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- For $p(S) = A \times S^B$, each state is assigned an element of A and a function $B \to S$ which updates the state
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

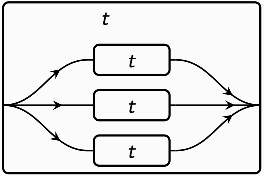
- A $[t^{\otimes m}, t]$-coalgebra consists of, for each state $s \in S$, a "wiring" $t^{\otimes m} \to t$ and a "rewiring" function $\mathbb{R}^m \times \mathbb{R} \to S$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- For $p(S) = A \times S^B$, each state is assigned an element of A and a function $B \to S$ which updates the state
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

- A $[t^{\otimes m}, t]$-coalgebra consists of, for each state $s \in S$, a "wiring" $t^{\otimes m} \to t$ and a "rewiring" function $\mathbb{R}^m \times \mathbb{R} \to S$



Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of “states” with a function $S \to p(S)$
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

- For deep learning:

$$S_{m,n} = \{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathsf{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathsf{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathsf{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

- For deep learning:

$$S_{m,n} = \{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$

- (k, f, r) is assigned $\mathbb{R}^m \xrightarrow{f(r,-)} \mathbb{R}^n$ and

$$\mathbb{R}^m \times \mathbb{R}^n \xrightarrow{\pi_2} \mathbb{R}^n \xrightarrow{Df^{\top}} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\pi_2} \mathbb{R}^m$$

Dynamics: polynomial coalgebras

- A p-coalgebra is a set S of "states" with a function $S \to p(S)$
- Let $t = \mathbb{R}y^{\mathbb{R}}$ and define the polynomial $[t^{\otimes m}, t^{\otimes n}]$ as

$$\mathrm{Hom}_{\mathbf{Poly}}(t^{\otimes m}, t^{\otimes n})y^{\mathbb{R}^m \times \mathbb{R}^n} = \mathrm{Hom}(\mathbb{R}^m, \mathbb{R}^n) \times \mathrm{Hom}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^m)y^{\mathbb{R}^m \times \mathbb{R}^n}$$

- For deep learning:

$$S_{m,n} = \{(k \in \mathbb{N}, f \colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$

- (k, f, r) is assigned $\mathbb{R}^m \xrightarrow{f(r,-)} \mathbb{R}^n$ and

$$\mathbb{R}^m \times \mathbb{R}^n \xrightarrow{\pi_2} \mathbb{R}^n \xrightarrow{Df^\top} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\pi_2} \mathbb{R}^m$$

- the feedback $x \in \mathbb{R}^n$ causes (k, f, r) to update to

$$(k, f, r + \epsilon \pi_1 Df^\top x)$$

Nesting: operad structure

Nesting: operad structure

- A *dynamic operad* on p is a sequence of coalgebras $S_n \to [p^{\otimes n}, p](S_n)$ for all n, along with coherent functions

$$1 \to S_1 \qquad \text{and} \qquad S_n \times S_{m_1} \times \cdots \times S_{m_n} \to S_{m_1+\cdots+m_n}$$

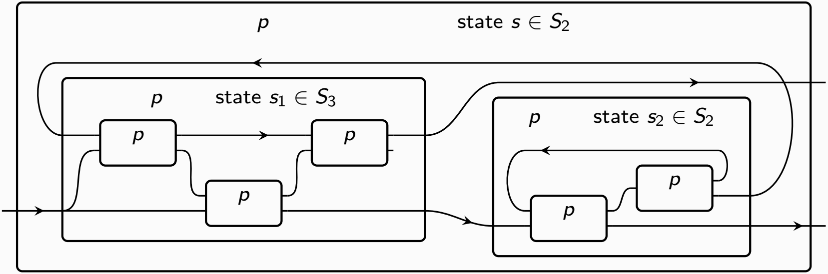
that respect identity and composition of morphisms into p

Nesting: operad structure

- A *dynamic operad* on p is a sequence of coalgebras $S_n \to [p^{\otimes n}, p](S_n)$ for all n, along with coherent functions

$$1 \to S_1 \qquad \text{and} \qquad S_n \times S_{m_1} \times \cdots \times S_{m_n} \to S_{m_1+\cdots+m_n}$$

that respect identity and composition of morphisms into p

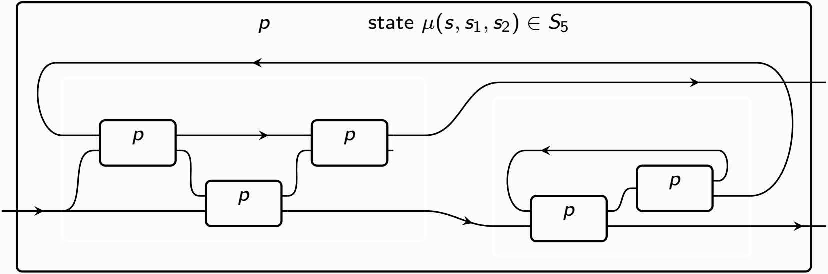


Nesting: operad structure

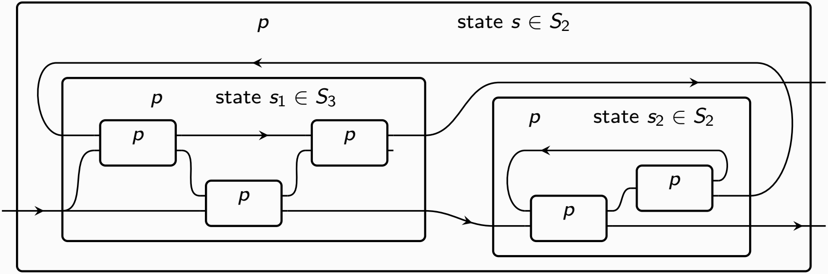
- A *dynamic operad* on p is a sequence of coalgebras $S_n \to [p^{\otimes n}, p](S_n)$ for all n, along with coherent functions

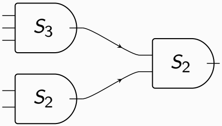
$$1 \to S_1 \qquad \text{and} \qquad S_n \times S_{m_1} \times \cdots \times S_{m_n} \to S_{m_1+\cdots+m_n}$$

that respect identity and composition of morphisms into p



- A *dynamic operad* on p is a sequence of coalgebras $S_n \to [p^{\otimes n}, p](S_n)$ for all n, along with coherent functions
$$1 \to S_1 \qquad \text{and} \qquad S_n \times S_{m_1} \times \cdots \times S_{m_n} \to S_{m_1+\cdots+m_n}$$
that respect identity and composition of morphisms into p



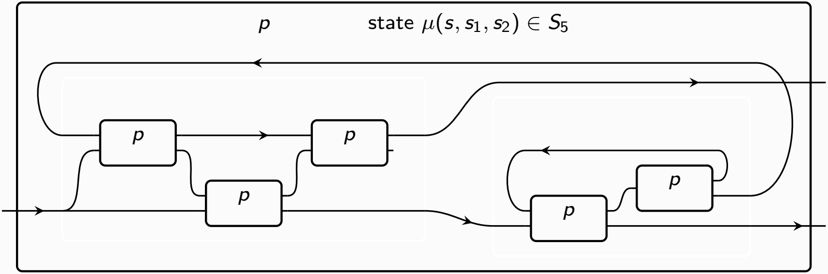


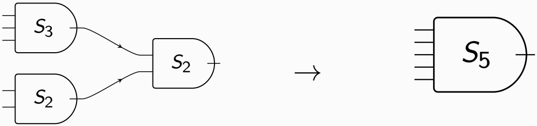
Nesting: operad structure

- A *dynamic operad* on p is a sequence of coalgebras $S_n \to [p^{\otimes n}, p](S_n)$ for all n, along with coherent functions

$$1 \to S_1 \qquad \text{and} \qquad S_n \times S_{m_1} \times \cdots \times S_{m_n} \to S_{m_1+\cdots+m_n}$$

that respect identity and composition of morphisms into p





Networks: PROP structure

Networks: PROP structure

- A *dynamic PROP* on p is a sequence of coalgebras $S_{m,n} \to [p^{\otimes m}, p^{\otimes n}](S_{m,n})$ for all m, n, along with functions

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

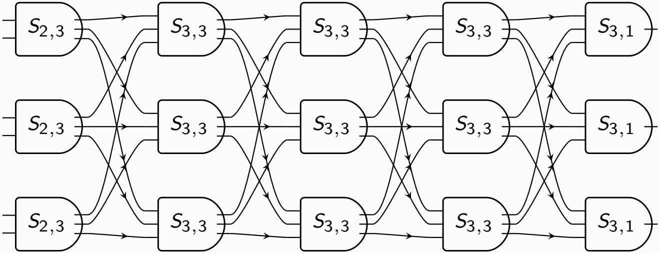
respecting identity/composition/tensor of morphisms into p

- A *dynamic PROP* on p is a sequence of coalgebras $S_{m,n} \to [p^{\otimes m}, p^{\otimes n}](S_{m,n})$ for all m, n, along with functions

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

respecting identity/composition/tensor of morphisms into p



A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

- $S_{m,n} =$

$$\{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$

A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

- $S_{m,n} =$

  $$\{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$

- $(0, \mathrm{id}_{\mathbb{R}}, *)$, $(0, \mathrm{id}_*, *)$

A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

- $S_{m,n} =$
$$\{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$
- $(0, \mathrm{id}_{\mathbb{R}}, *)$, $(0, \mathrm{id}_*, *)$
- $(\sigma, \tau, (k, f, r)) \mapsto$
$$(k, \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\mathrm{id}\times\mathbb{R}^\sigma} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{\mathbb{R}^\tau} \mathbb{R}^n, r)$$

A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

- $S_{m,n} =$
$$\{(k \in \mathbb{N}, f \colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$
- $(0, \mathrm{id}_{\mathbb{R}}, *)$, $(0, \mathrm{id}_*, *)$
- $(\sigma, \tau, (k, f, r)) \mapsto$
$$(k, \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\mathrm{id} \times \mathbb{R}^\sigma} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{\mathbb{R}^\tau} \mathbb{R}^n, r)$$
- $((k', f', r'), (k, f, r)) \mapsto$
$$(k + k', \mathbb{R}^k \times \mathbb{R}^{k'} \times \mathbb{R}^\ell \xrightarrow{\mathrm{id} \times f'} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{f} \mathbb{R}^n, (r, r'))$$

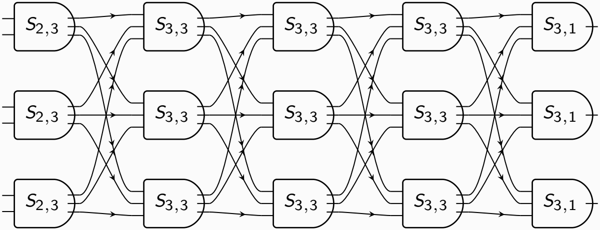
A dynamic PROP for deep learning

$$1 \to S_{1,1}, \quad 1 \to S_{0,0}, \quad \mathbf{Perm}(m) \times \mathbf{Perm}(n) \times S_{m,n} \to S_{m,n}$$

$$S_{\ell,m} \times S_{m,n} \to S_{\ell,n}, \quad \text{and} \quad S_{m_1,n_1} \times S_{m_2,n_2} \to S_{m_1+m_2,n_1+n_2}$$

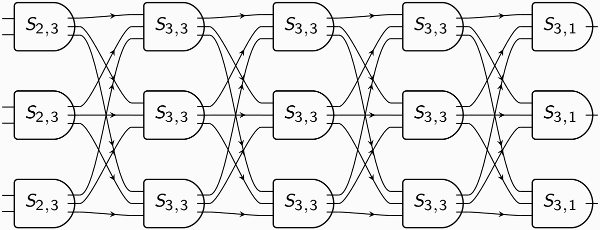
- $S_{m,n} =$
$$\{(k \in \mathbb{N}, f\colon \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^n, r \in \mathbb{R}^k) | f \text{ is differentiable}\}$$
- $(0, \mathrm{id}_{\mathbb{R}}, *)$, $(0, \mathrm{id}_*, *)$
- $(\sigma, \tau, (k, f, r)) \mapsto$
$$(k, \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{\mathrm{id}\times\mathbb{R}^\sigma} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{\mathbb{R}^\tau} \mathbb{R}^n, r)$$
- $((k', f', r'), (k, f, r)) \mapsto$
$$(k + k', \mathbb{R}^k \times \mathbb{R}^{k'} \times \mathbb{R}^\ell \xrightarrow{\mathrm{id}\times f'} \mathbb{R}^k \times \mathbb{R}^m \xrightarrow{f} \mathbb{R}^n, (r, r'))$$
- $((k_1, f_1, r_1), (k_2, f_2, r_2)) \mapsto (k_1 + k_2, f_1 \times f_2, (r_1, r_2))$

Vistas



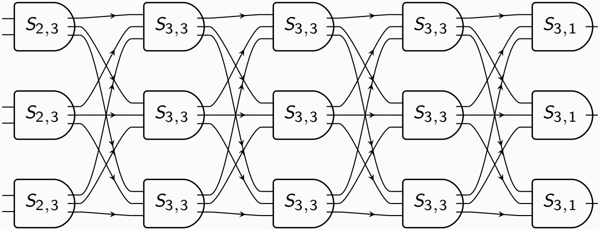
Vistas

- Model other types of learning systems compositionally as dynamic categorical structures



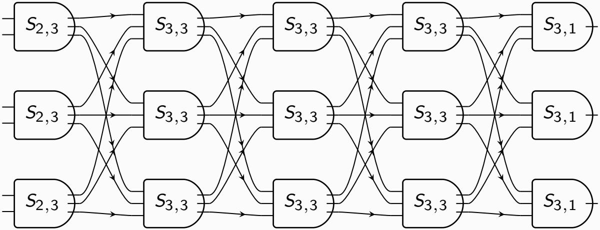
Vistas

- Model other types of learning systems compositionally as dynamic categorical structures
- As a well behaved categorical structure, dynamic PROPS can be easily compared, combined, or generalized



Vistas

- Model other types of learning systems compositionally as dynamic categorical structures
- As a well behaved categorical structure, dynamic PROPS can be easily compared, combined, or generalized
- Implementation in computing frameworks based on category theory or polynomials (algebraicjulia.org)



References

- Brandon T. Shapiro and David I. Spivak, "Dynamic operads, dynamic categories: From deep learning to prediction markets" arXiv:2205.03906
- David I. Spivak, "Learners' Languages" arXiv:2103.01189
- Sophie Libkind and David I. Spivak, "When you light up, I light up: A dynamical monoidal category of Hebbian learners" Topos Institute Blog

Thanks!