# Softmax is Argmax, and the Logic of the Reals

MSP101

Matteo Capucci

2024-03-22

University of Strathclyde Glasgow

# Prologue

One of me toxic traits is pointing at things and *justify* them: ‘that’s a coend’, ‘that’s a cartesian lift’, ‘that’s an enriched quasi-limit’.

[1] Debatable.

# Prologue

One of me toxic traits is pointing at things and *justify* them: ‘that’s a coend’, ‘that’s a cartesian lift’, ‘that’s an enriched quasi-limit’.

One could call this ‘semiotics’[1]: analyzing a **sign** to understand how it decouples in a **signifier** (a syntactic hence logical/categorical object) with a certain **semantics** (a choice of ambient to interpret it).

[1] Debatable.

# Prologue

One of me toxic traits is pointing at things and *justify* them: ‘that’s a coend’, ‘that’s a cartesian lift’, ‘that’s an enriched quasi-limit’.

One could call this ‘semiotics’[1]: analyzing a **sign** to understand how it decouples in a **signifier** (a syntactic hence logical/categorical object) with a certain **semantics** (a choice of ambient to interpret it).

For a while I’ve been thinking about argmax and softmax: can we make sense of their similarity?

[1] Debatable.

# Softmax

**Definition**. Let $f : X \to [0, \infty]$ be an integrable function over $X$ measure space. Its **softmax** is the probability distribution $X \to [0, 1]$ defined as

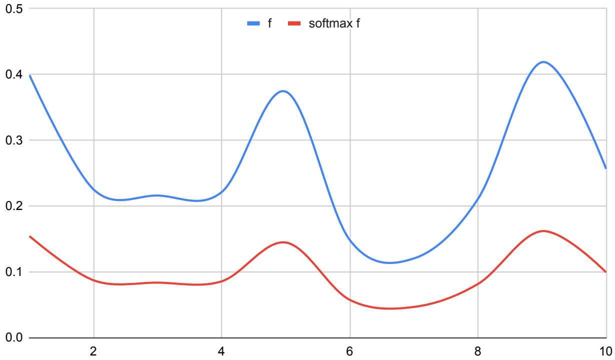
$$(\mathrm{softmax}\, f)(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx}.$$

# Softmax

**Definition.** Let $f : X \to [0, \infty]$ be an integrable function over $X$ measure space. Its **softmax** is the probability distribution $X \to [0, 1]$ defined as

$$(\mathrm{softmax}\, f)(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx}.$$

1. not as brittle than argmax
2. all the cool kids are doing it



# Softmax vs. argmax

**Definition**. Let $f : X \to [0, \infty]$ be an integrable function over $X$ measure space. Its **softmax** is the probability distribution $X \to [0, 1]$ defined as

$$(\mathrm{softmax}\, f)(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx}.$$

**Definition**. Let $f : X \to \mathbb{R}$ be a function. Its **argmax** is the predicate $X \to \mathbf{Prop}$ defined as

$$(\mathrm{argmax}\, f)(x^*) = \forall x \in X,\ f(x) \le f(x^*).$$

# Softmax vs. argmax

**Definition**. Let $f : X \to [0, \infty]$ be an integrable function over $X$ measure space. Its **softmax** is the probability distribution $X \to [0, 1]$ defined as

$$(\operatorname{softmax} f)(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx}.$$

**Definition**. Let $f : X \to \mathbb{R}$ be a function. Its **argmax** is the predicate $X \to$ **Prop** defined as

$$(\operatorname{argmax} f)(x^*) = \forall x \in X,\ f(x) \leq f(x^*).$$

# Idea

In general, the fact that $\int$ 'behaves like' '$\exists$/*colim*' has been noted before:

From (Loregian 2021):

$$\int_{x\in\mathbb{R}} f(x) \circ \delta(x-y)dx = f(y)$$

$$\int^{X} FX \times \mathcal{C}(Y,X) \cong FY$$

The analogy between the pairing of a function and a delta distribution, and the ninja Yoneda lemma.

# Idea

In general, the fact that $\int$ 'behaves like' '$\exists$/*colim*' has been noted before:

From (Perrone and Tholen 2021):

- small presheaves on a category are similar to measures on a measurable space;
- cocomplete categories (categories where one can take colimits) are similar to algebras over probability monads (spaces where one can take integrals or expectation values, such as the real line);

# Idea

1. Can the analogy be made precise?

# Idea

1. Can the analogy be made precise?
2. What are even quantifiers in this setting?

# Idea

1. Can the analogy be made precise?
2. What are even quantifiers in this setting?

One could say I'm trying to combine two of Lawvere's most famous papers:

- 'Adjointess in Foundations' (Lawvere 1969) and
- 'Metric spaces, generalized logic, and closed categories' (Lawvere 1973).

# Idea

1. Can the analogy be made precise?
2. What are even quantifiers in this setting?

One could say I’m trying to combine two of Lawvere’s most famous papers:

- ‘Adjointess in Foundations’ (Lawvere 1969) and
- ‘Metric spaces, generalized logic, and closed categories’ (Lawvere 1973).

Radu & collaborators have done lots of work in this direction already! I’m an amateur.

# Logic of the Reals

# The multiplicative reals

**Definition.** The poset $([0, \infty], \leq)$ is a $*$-autonomous quantale, meaning:

1. it has all small joins $\bigvee_i a_i$, given by suprema;

# The multiplicative reals

**Definition.** The poset $([0, \infty], \leq)$ is a $*$-autonomous quantale, meaning:

1. it has all small joins $\bigvee_i a_i$, given by suprema;
2. it is equipped with a join-preserving tensor product $\otimes : [0, \infty] \times [0, \infty] \to [0, \infty]$, defined by multiplication and satisfying commutativity, associativity and unitality, with unit $1 \in [0, \infty]$.

# The multiplicative reals

**Definition**. The poset $([0,\infty],\leq)$ is a $*$-autonomous quantale, meaning:

1. it has all small joins $\bigvee_i a_i$, given by suprema;
2. it is equipped with a join-preserving tensor product $\otimes : [0,\infty] \times [0,\infty] \to [0,\infty]$, defined by multiplication and satisfying commutativity, associativity and unitality, with unit $1 \in [0,\infty]$.
3. it is equipped with a duality $(-)^* : [0,\infty]^{\mathrm{op}} \to [0,\infty]$, defined as

$$\forall a \in (0,\infty),\ a^* := 1/a, \quad 0^* = \infty,\ \infty^* = 0,$$

   satisfying the property

$$\forall a \in [0,\infty],\ a \otimes b \leq c^* \iff a \leq (b \otimes c)^*.$$

# $0\infty = ?$

The definition of $\otimes$ involving $\infty$ is dictated by the requirement $\otimes$ preserves joins: for any unbounded set $(u_i)_{i \in I}$, we have

$$a \otimes \infty := a \otimes \bigvee_i u_i = \bigvee_i a \otimes u_i = \infty$$

$$0 \otimes \infty := 0 \otimes \bigvee_i u_i = \bigvee_i 0 \otimes u_i = 0$$

# $0\infty = ?$

The definition of $\otimes$ involving $\infty$ is dictated by the requirement $\otimes$ preserves joins: for any unbounded set $(u_i)_{i \in I}$, we have

$$a \otimes \infty := a \otimes \bigvee_i u_i = \bigvee_i a \otimes u_i = \infty$$

$$0 \otimes \infty := 0 \otimes \bigvee_i u_i = \bigvee_i 0 \otimes u_i = 0$$

Also: the law $(a \otimes b)^* = a^* \otimes b^*$ isn't required to hold. In fact

$$0 = 0 \otimes \infty = \infty^* \otimes 0^* \neq (\infty \otimes 0)^* = \infty$$

It does hold for all $0 < a, b < \infty$ though.

# $0\infty = ?$

The reason this doesn't hold is that a $*$-**autonomous category** really features two distinct monoidal products, where the second is de Morgan dual to the first:

$$a \otimes^* b = (a^* \otimes b^*)^*$$

In the case of $[0, \infty]$, $\otimes^*$ coincides with $\otimes$ except for the rule

$$0 \otimes^* \infty = \infty$$

# $0\infty = ?$

The reason this doesn't hold is that a $*$**-autonomous category** really features two distinct monoidal products, where the second is de Morgan dual to the first:

$$a \otimes^* b = (a^* \otimes b^*)^*$$

In the case of $[0, \infty]$, $\otimes^*$ coincides with $\otimes$ except for the rule

$$0 \otimes^* \infty = \infty$$

Thus they (lax) linearly distribute over each other:

$$a \otimes (b \otimes^* c) \leq (a \otimes b) \otimes^* c$$

where, again, equality holds for all finite values and breaks for $(a, b, c) = (0, 0, \infty)$.

# Closure

**Bonus**: every $*$-autonomous category is closed (not compact though):

$$a \multimap b := a^* \otimes b = \begin{cases} 0 & b = 0 \text{ or } a = \infty \\ \infty & a = 0 \text{ and } b \neq \infty \\ b/a & \text{else} \end{cases}$$

# Closure

**Bonus**: every $*$-autonomous category is closed (not compact though):

$$a \multimap b := a^* \otimes b = \begin{cases} 0 & b = 0 \text{ or } a = \infty \\ \infty & a = 0 \text{ and } b \neq \infty \\ b/a & \text{else} \end{cases}$$

...and coclosed:

$$a \bullet\!\!- b := a \otimes^* b^* = \begin{cases} \infty & a = 0 \text{ or } b = \infty \\ 0 & b = 0 \text{ and } a \neq \infty \\ a/b & \text{else} \end{cases}$$

# Sums

Crucially, $[0, \infty]$ also has sums:

$$a \oplus b := \begin{cases} \infty & a = \infty \text{ or } b = \infty \\ a + b & \text{else} \end{cases}$$

which are commutative, associative and unital, with unit 0;

# Sums

Crucially, $[0, \infty]$ also has sums:

$$a \oplus b := \begin{cases} \infty & a = \infty \text{ or } b = \infty \\ a + b & \text{else} \end{cases}$$

which are commutative, associative and unital, with unit 0;

and **harmonic sums**:

$$a \oplus^* b := (a^* \oplus b^*)^* = \begin{cases} 0 & a = 0 \text{ or } b = 0 \\ a & b = \infty \\ b & a = \infty \\ \dfrac{1}{1/a + 1/b} & \text{else} \end{cases}$$

which are commutative, associative and unital, with unit $\infty$.

# Algebro-logical structures on the multiplicative reals

| | | | linear | |
|---|---|---|---|---|
| polarity | | non-linear | additive | multiplicative |
| **duality** $a^* := 1/a$ | **positive** | **false** $:= 0$ <br> $a \vee b := \max\{a, b\}$ | $\mathbf{0} := 0$ <br> $a \oplus b := a + b$ | $\mathbf{1} := 1$ <br> $a \otimes b := ab,\ 0\infty = 0$ |
| | **negative** | **true** $:= \infty$ <br> $a \wedge b := \min\{a, b\}$ | $\top := \infty$ <br> $a \oplus^* b := a \oplus^* b$ | $\bot := 1$ <br> $a \otimes^* b := ab,\ 0\infty = \infty$ |

# Algebro-logical structures on the multiplicative reals

| | | non-linear | linear | |
|---|---|---|---|---|
| polarity | | | additive | multiplicative |
| duality $a^* := 1/a$ | positive | $\textbf{false} := 0$<br>$a \vee b := \max\{a, b\}$ | $\mathbf{0} := 0$<br>$a \oplus b := a + b$ | $\mathbf{1} := 1$<br>$a \otimes b := ab,\ 0\infty = 0$ |
| | negative | $\textbf{true} := \infty$<br>$a \wedge b := \min\{a, b\}$ | $\top := \infty$<br>$a \oplus^* b := a \oplus^* b$ | $\bot := 1$<br>$a \otimes^* b := ab,\ 0\infty = \infty$ |

and three kinds of distributivity (+ duals not shown):

**multiplicative** $a \otimes^* (b \otimes c) \leq (a \otimes^* b) \otimes c$ $\quad \mathbf{1} \leq \top$

**multiplicative-additive** $a \otimes (b \oplus_p c) = (a \otimes b) \oplus_p (a \otimes c)$ $\quad a \otimes 0 = 0$

**linear-nonlinear** $a \otimes (b \vee c) = (a \otimes b) \vee (a \otimes c)$ $\quad a \otimes \textbf{false} = \textbf{false}$

$a \oplus (b \vee c) = (a \oplus b) \vee (a \oplus c)$ $\quad a \oplus \textbf{true} = \textbf{true}$

# Conjectures

I'm not a logician so don't laugh:

**Conjecture 1.** There is an extension of classical linear logic featuring an extra generation of linear connectives:

| linear | 'very linear' |
|---|---|
| **additive** $\vee, \wedge$ | **non-linear** $\vee, \wedge$ |
| **multiplicative** $\otimes, \otimes^*$ | **linear additive** $\oplus, \oplus^*$<br>**linear multiplicative** $\otimes, \otimes^*$ |

*Open question: what are the rules associated to $\oplus$ and $\oplus^*$?*

# Conjectures

**Conjecture 2.** The Lambek side of this logic is

**$*$-autonomous symmetric bimonoidal bicartesian categories.**

A **symmetric bimonoidal category** is a category $\mathbf{A}$ equipped with two monoidal structures $(\mathbf{1}, \otimes)$ and $(\mathbf{0}, \oplus)$ and coherent isomorphisms:

$$a \otimes (b \oplus c) \cong (a \otimes b) \oplus (a \otimes c), \qquad a \otimes \mathbf{0} \cong \mathbf{0}.$$

It is **$*$-autonomous** when it is a Frobenius pseudomonoid, i.e. it comes with a symmetric monoidal functor $(-)^* : \mathbf{A}^{\mathrm{op}} \longrightarrow \mathbf{A}$ and natural isomorphisms

$$\mathbf{A}(a \otimes b, c^*) \cong \mathbf{A}(a, (b \otimes c)^*).$$

It is **bicartesian** when instead of starting from a category $\mathbf{A}$, we start from a bicartesian category $\mathbf{A}$ (so that all operations commute with $\bigvee$).

# Comparison with Polynomial Lawvere Logic ($\mathbb{PL}$)

Recently Radu & co[2] introduced $\mathbb{PL}$, a logic valued in $([0,\infty],\geq,\oplus)$ with $\otimes$ (my notation).

The link with the above is given by the following well-known diagram of quantales:

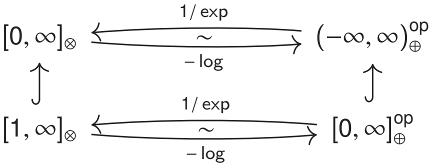
$$
\begin{array}{ccc}
[0,\infty]_{\otimes} & \overset{1/\exp}{\underset{-\log}{\rightleftarrows}} \; \sim & (-\infty,\infty)^{\mathrm{op}}_{\oplus} \\
\uparrow & & \uparrow \\
[1,\infty]_{\otimes} & \overset{1/\exp}{\underset{-\log}{\rightleftarrows}} \; \sim & [0,\infty]^{\mathrm{op}}_{\oplus}
\end{array}
$$

[2] (Bacci, Mardare, Panangaden, and Plotkin 2024)

# Comparison with Polynomial Lawvere Logic ($\mathbb{PL}$)

Recently Radu & co[2] introduced $\mathbb{PL}$, a logic valued in $([0,\infty], \geq, \oplus)$ with $\otimes$ (my notation).

The link with the above is given by the following well-known diagram of quantales:




Some observations:

1. the $\otimes$ world seems to 'have more natural structure': $1/\exp$ and $-\log$ don't see $\oplus$ on the left
2. working with unsigned truth values allows to represent more 'degrees of falsehood'
3. additive enrichment seems to work better than multiplicative one
4. less obvious candidates for quantifiers: can't use $\oplus$ for $\exists$.

[2](Bacci, Mardare, Panangaden, and Plotkin 2024)

# Coda: from quantitative to qualitative

To extract ‘qualitative’ truth from ‘quantitative’ one chooses an **ideal** representing truth:

$$F \subseteq [0,\infty]_\otimes \text{ which is } \begin{cases} \text{inhabited} & \infty \in F, \\ \text{upper-closed} & a \in F, a \leq b \implies b \in F \\ \otimes\text{-closed} & a, b \in F \implies a \otimes b \in F \end{cases}$$

# Coda: from quantitative to qualitative

To extract 'qualitative' truth from 'quantitative' one chooses an **ideal** representing truth:

$$F \subseteq [0,\infty]_\otimes \text{ which is } \begin{cases} \text{inhabited} & \infty \in F, \\ \text{upper-closed} & a \in F, a \leq b \implies b \in F \\ \otimes\text{-closed} & a, b \in F \implies a \otimes b \in F \end{cases}$$

**Lemma**. Every ideal in $[0,\infty]_\otimes$ is principal and has one of four forms:

$$\{\infty\} \qquad [1+\varepsilon,\infty], \varepsilon \geq 0 \qquad (0,\infty] \qquad [0,\infty]$$

# Coda: from quantitative to qualitative

To extract ‘qualitative’ truth from ‘quantitative’ one chooses an **ideal** representing truth:

$$F \subseteq [0,\infty]_\otimes \text{ which is } \begin{cases} \text{inhabited} & \infty \in F, \\ \text{upper-closed} & a \in F, a \leq b \implies b \in F \\ \otimes\text{-closed} & a, b \in F \implies a \otimes b \in F \end{cases}$$

**Lemma**. Every ideal in $[0,\infty]_\otimes$ is principal and has one of four forms:

$$\{\infty\} \qquad [1+\varepsilon,\infty], \varepsilon \geq 0 \qquad (0,\infty] \qquad [0,\infty]$$

**Definition**. We say $\varphi \in [0,\infty]_\otimes$ is ***F*-qualitatively true** when $\varphi \in F$.

# Coda: from quantitative to qualitative

To extract ‘qualitative’ truth from ‘quantitative’ one chooses an **ideal** representing truth:

$$F \subseteq [0,\infty]_\otimes \ \text{ which is } \begin{cases} \text{inhabited} & \infty \in F, \\ \text{upper-closed} & a \in F, a \le b \implies b \in F \\ \otimes\text{-closed} & a, b \in F \implies a \otimes b \in F \end{cases}$$

**Lemma**. Every ideal in $[0,\infty]_\otimes$ is principal and has one of four forms:

$$\{\infty\} \qquad [1+\varepsilon,\infty], \varepsilon \ge 0 \qquad (0,\infty] \qquad [0,\infty]$$

**Definition**. We say $\varphi \in [0,\infty]_\otimes$ is ***F*-qualitatively true** when $\varphi \in F$.

My favourite ideal is $[1,\infty]$, and it corresponds to ‘positive evidence’. Indeed this the choice made in the additive world of $\mathbb{PL}$.

# Quantifiers

# The spectrum of $p$-sums

**Definition**. For any $p \in (-\infty, \infty)$, $p \neq 0$, the ***p*-sum** of a finite set of numbers $(a_i)_{i \in I}$ is

$$\bigoplus_{i \in I}^{p} a_i := \left( \bigoplus_{i \in I} a_i^p \right)^{1/p}$$

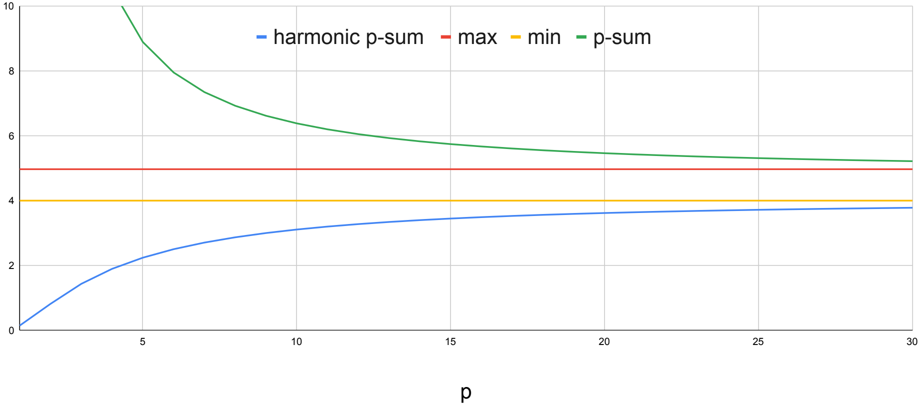
when $p < 0$ we call $p$-sums **harmonic**.

One extends the above definition to $p = \pm\infty$ (but not $p = 0$) by taking suitable limits.

$$\begin{array}{ccccccccccccc} \bigwedge & \cdots & \bigoplus^{-p} & \cdots & \bigoplus^{*} & \infty & \bigoplus & \cdots & \bigoplus^{p} & \cdots & \bigvee \\ -\infty & \text{———} & -p & \text{———} & -1 & \text{——} & 0 & \text{——} & 1 & \text{———} & p & \longrightarrow & \infty \end{array}$$

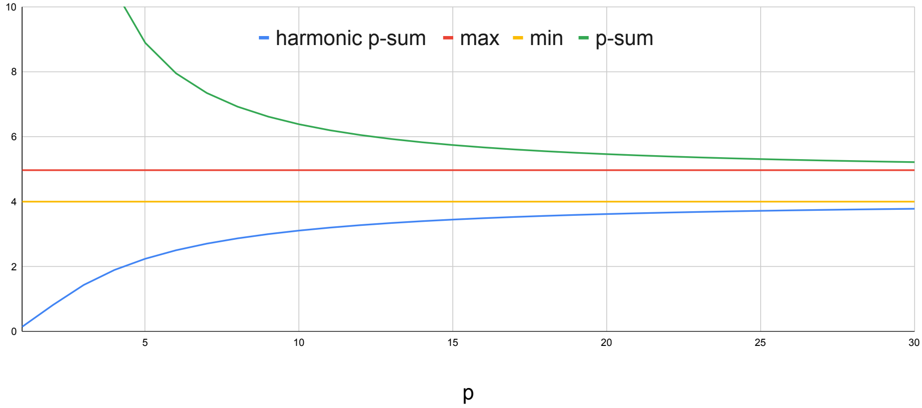
# The spectrum of $p$-sums

We have $\bigoplus^{-p} \leq \bigwedge \leq \bigvee \leq \bigoplus^{p}$, with the (exterior) gap narrowing as $p$ increases:



# The spectrum of $p$-sums

We have $\bigoplus^{-p} \leq \bigwedge \leq \bigvee \leq \bigoplus^{p}$, with the (exterior) gap narrowing as $p$ increases:




So deforming $\oplus$ to $\oplus^p$ doesn't change the 'logic' of $[0, \infty]_\otimes$, and in fact it converges to classical linear logic as $p \to \infty$.

# The amazing $p$-sums

The **fundamental relation of harmonic sums** is:

$$\bigoplus_{i\in I}^{-p}(b_i \multimap a) = \left(\bigoplus_{i\in I}^{p} b_i\right) \multimap a \quad \rightsquigarrow \bigoplus_{i\in I}^{-p}\frac{a}{b_i} = \frac{a}{\bigoplus_{i\in I}^{p} b_i}$$

which is analogous to $\forall i\,(b_i \to a) = (\exists i\, b_i) \to a$.

# The amazing $p$-sums

The **fundamental relation of harmonic sums** is:

$$\bigoplus_{i\in I}^{-p}(b_i \multimap a) = \left(\bigoplus_{i\in I}^{p} b_i\right) \multimap a \quad \rightsquigarrow \bigoplus_{i\in I}^{-p} \frac{a}{b_i} = \frac{a}{\bigoplus_{i\in I}^{p} b_i}$$

which is analogous to $\forall i\,(b_i \to a) = (\exists i\, b_i) \to a$.

In fact we also have:

$$\bigoplus_{i}^{-p}(b \multimap a_i) = b \multimap \left(\bigoplus_{i}^{-p} a_i\right),$$

which is analogous to $\forall i\,(b \to a_i) = b \to (\forall i\, a_i)$.

# The amazing $p$-sums

**Lemma**. Sum and harmonic sums...

1. are monotonic in the argument: if, for each $i \in I$, $a_i \leq b_i$, then

$$\bigoplus_{i \in I}^{*} a_i \leq \bigoplus_{i \in I}^{*} b_i,$$

# The amazing $p$-sums

**Lemma**. Sum and harmonic sums...

1. are monotonic in the argument: if, for each $i \in I$, $a_i \leq b_i$, then

$$\bigoplus_{i \in I}^{*} a_i \leq \bigoplus_{i \in I}^{*} b_i,$$

2. are, resp. monotonic and antitonic in the index: when $J \subseteq I$, one has

$$\bigoplus_{i \in I}^{p} a_i \geq \bigoplus_{j \in J}^{p} a_j, \qquad \bigoplus_{i \in I}^{-p} a_i \leq \bigoplus_{j \in J}^{-p} a_j.$$

# The amazing $p$-sums

**Lemma**. Sum and harmonic sums...

1. are monotonic in the argument: if, for each $i \in I$, $a_i \leq b_i$, then

$$\bigoplus_{i\in I}^{*} a_i \leq \bigoplus_{i\in I}^{*} b_i,$$

2. are, resp. monotonic and antitonic in the index: when $J \subseteq I$, one has

$$\bigoplus_{i\in I}^{p} a_i \geq \bigoplus_{j\in J}^{p} a_j, \qquad \bigoplus_{i\in I}^{-p} a_i \leq \bigoplus_{j\in J}^{-p} a_j.$$

However... ***p*-sums are susceptible to cumulative effects**: their truth value can pass any threshold given enough 'false' values (too easy to satisfy); *vice versa* for harmonic $p$-sums (too hard to satisfy). **We need to compensate for size!**

# The spectrum of $p$-means

**Definition**. For any $p \in (-\infty, \infty)$, $p \neq 0$, the ***p*-mean** of a finite set of numbers $(a_i)_{i \in I}$ is

$$\int_{i \in I}^{p} a_i := \left( \bigoplus_{i \in I}^{p} \frac{a_i^p}{|I|} \right)^{1/p}$$

when $p < 0$ we call $p$-means **harmonic.**

One extends the above definition to $p = \pm\infty$ by taking suitable limits.

$$\begin{array}{ccccccccccccc} \bigwedge & \cdots & \int^{-p} & \cdots & \int^{*} & GM & \int & \cdots & \int^{p} & \cdots & \bigvee \\ -\infty & \text{———} & -p & \text{———} & -1 & \text{——} & 0 & \text{——} & 1 & \text{———} & p & \longrightarrow & \infty \end{array}$$

# The spectrum of $p$-means

**Definition**. For any $p \in (-\infty, \infty)$, $p \neq 0$, the ***p*-mean** of a finite set of numbers $(a_i)_{i \in I}$ is

$$\int_{i \in I}^{p} a_i := \left( \bigoplus_{i \in I}^{p} \frac{a_i^p}{|I|} \right)^{1/p}$$

when $p < 0$ we call $p$-means **harmonic.**

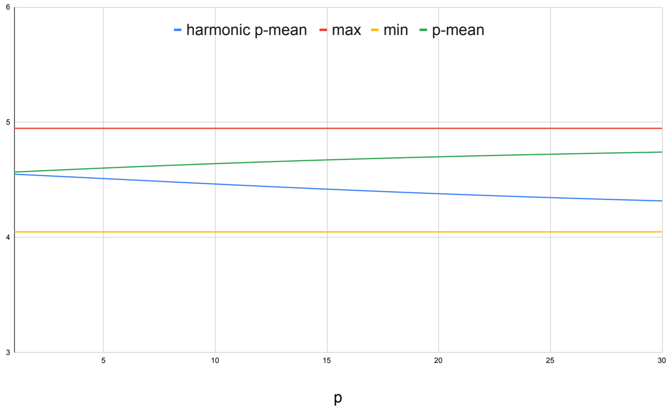
One extends the above definition to $p = \pm\infty$ by taking suitable limits.

$$\begin{array}{ccccccccccccc} \bigwedge & \cdots & \int^{-p} & \cdots & \int^{*} & GM & \int & \cdots & \int^{p} & \cdots & \bigvee \\ -\infty & \text{———} & -p & \text{———} & -1 \text{ —— } & 0 & \text{—— } 1 & \text{———} & p & \longrightarrow & \infty \end{array}$$

**Fact**. $p$-means satisfy the same properties as $p$-sums: Fubini, monotonicity in the argument and index, fundamental relation and homogeneity.

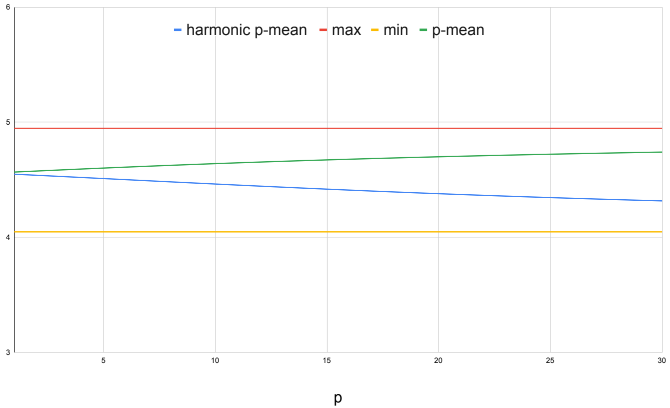
# The spectrum of $p$-means

We have $\bigwedge \leq \int^{-p} \leq \int^{p} \leq \bigvee$, with the (exterior) gap narrowing as $p$ increases:



# The spectrum of $p$-means

We have $\bigwedge \leq \int^{-p} \leq \int^{p} \leq \bigvee$, with the (exterior) gap narrowing as $p$ increases:




Unlike $p$-sums, $p$-means **compensate for cumulative effects.**

# Means as bounded quantifiers

**Idea.** Mean and harmonic mean correspond to **bounded quantification**:

$$\exists i.\ (i \in I) \wedge a(i) \quad \leftrightsquigarrow \quad \bigoplus_{i \in I}^{p} \frac{1}{|I|} \otimes a(i)$$

# Means as bounded quantifiers

**Idea.** Mean and harmonic mean correspond to **bounded quantification**:

$$\exists i.\ (i \in I) \wedge a(i) \quad \leftrightsquigarrow \quad \bigoplus_{i \in I}^{p} \frac{1}{|I|} \otimes a(i)$$

A mean is a just an integral over a probability space so we can directly generalize:

$$\exists i \in I.\ a(i) \quad \leftrightsquigarrow \quad \int_{i \in I}^{p} a(i)\, di.$$

for $I$ any probability space.

# Means as bounded quantifiers

**Idea.** Mean and harmonic mean correspond to **bounded quantification**:

$$\exists i.\ (i \in I) \wedge a(i) \quad \leftrightsquigarrow \quad \bigoplus_{i \in I}^{p} \frac{1}{|I|} \otimes a(i)$$

A mean is a just an integral over a probability space so we can directly generalize:

$$\exists i \in I.\ a(i) \quad \leftrightsquigarrow \quad \int_{i \in I}^{p} a(i)\, di.$$

for $I$ any probability space. Duality also suggests this interpretation since

$$\forall i \in I.\ a(i) = \neg \exists i \in I.\ \neg a(i) \quad \leftrightsquigarrow \quad \int_{i \in I}^{-p} a(i) = \left( \int_{i \in I}^{p} a(i)^{*}\, di \right)^{*},$$

i.e. we dualize $a(i)$ but not the domain of quantification.

# Softmax is argmax!

Given the premise, it's not hard to see what we are going to propose.

# Softmax is argmax!

Given the premise, it's not hard to see what we are going to propose. For $X$ probability space and $f : X \longrightarrow [0, \infty]$, define

$$(p\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{-p} f(x) \multimap f(x^*).$$

# Softmax is argmax!

Given the premise, it's not hard to see what we are going to propose. For $X$ probability space and $f : X \longrightarrow [0, \infty]$, define

$$(p\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{-p} f(x) \multimap f(x^*).$$

Then it's easy to see that $\operatorname{softmax} = 1\text{-}\operatorname{softmax}$:

$$(1\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{*} f(x) \multimap f(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx} = (\operatorname{softmax} f)(x^*)$$

# Softmax is argmax!

Given the premise, it's not hard to see what we are going to propose. For $X$ probability space and $f : X \longrightarrow [0, \infty]$, define

$$(p\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{-p} f(x) \multimap f(x^*).$$

Then it's easy to see that $\operatorname{softmax} = 1\text{-}\operatorname{softmax}$:

$$(1\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{*} f(x) \multimap f(x^*) = \frac{f(x^*)}{\int_{x \in X} f(x)\, dx} = (\operatorname{softmax} f)(x^*)$$

Conversely, $\operatorname{argmax}$ is $[1, \infty]$-qualitatively $\infty\text{-}\operatorname{softmax}$:

$$(\infty\text{-}\operatorname{softmax} f)(x^*) = \int_{x \in X}^{-\infty} f(x) \multimap f(x^*) = \bigwedge_{x \in X} \frac{f(x^*)}{f(x)} \overset{\lfloor - \rfloor}{\longmapsto} (\operatorname{argmax} f)(x^*)$$

The passage to qualitative is necessary since that's where $\operatorname{argmax}$ lives!

# Conjecture: transporting quantifiers to $\mathbb{PL}$

One can transport structure from $[0, \infty]_{\otimes}$ to $[-\infty, \infty]_{\oplus}$ along Napier's isomorphism.

Thus for $\varphi : I \to [-\infty, \infty]_{\oplus}$, define:

$$\exists^p (i \in I).\ \varphi(i) := -\log \int_{i \in I}^{p} \exp(-\varphi(i))\, di$$

$$\forall^p (i \in I).\ \varphi(i) := -\log \int_{i \in I}^{-p} \exp(-\varphi(i))\, di = \log \int_{i \in I}^{p} \exp(\varphi(i))\, di$$

# Conjecture: transporting quantifiers to $\mathbb{PL}$

One can transport structure from $[0, \infty]_\otimes$ to $[-\infty, \infty]_\oplus$ along Napier's isomorphism.

Thus for $\varphi : I \to [-\infty, \infty]_\oplus$, define:

$$\exists^p (i \in I).\ \varphi(i) := -\log \int_{i \in I}^{p} \exp(-\varphi(i))\, di$$

$$\forall^p (i \in I).\ \varphi(i) := -\log \int_{i \in I}^{-p} \exp(-\varphi(i))\, di = \log \int_{i \in I}^{p} \exp(\varphi(i))\, di$$

It's also very typical to start with $u : X \to [-\infty, \infty]_\oplus$ and then construct $(\mathrm{softmax}\, e^{-u})(x^*)$. In logistic regression, one actually goes all the way back to obtain a quantity called **log-likelihood**:

$$L_u = -\log \mathrm{softmax}\, e^{-u} = u - \log \int_{x \in X} e^{-u(x)}\, dx = \text{``}\forall^p (x \in X).\ u(x) \multimap u(x^*)\text{''}$$

And this is indeed $\mathrm{argmax}\, u$ in $[-\infty, \infty]_\oplus$ according to the proposed quantifiers.

# Exploration:
# generalized logic via enriched category theory

# The mirage of an enriched hyperdoctrine

To each probability space $I$ we can try to associate an $[0,\infty]_\otimes$-enriched $p$-**Lindenbaum–Tarski algebra** of real-valued predicates:

$$\mathbf{LT}^p(I,[0,\infty]) = \begin{cases} \text{elements} & \varphi : I \to [0,\infty]_\otimes \\ \text{entailment} & \varphi \vdash_I \psi := \int_{i\in I}^{-p} \varphi(i) \multimap \psi(i) = \left(\int_{i\in I} \frac{\varphi(i)^p}{\psi(i)^p}\, di\right)^{-1/p} \end{cases}$$

# The mirage of an enriched hyperdoctrine

To each probability space $I$ we can try to associate an $[0,\infty]_\otimes$-enriched $p$**-Lindenbaum–Tarski algebra** of real-valued predicates:

$$\mathbf{LT}^p(I,[0,\infty]) = \begin{cases} \text{elements} & \varphi : I \to [0,\infty]_\otimes \\ \text{entailment} & \varphi \vdash_I \psi := \int_{i\in I}^{-p} \varphi(i) \multimap \psi(i) = \left(\int_{i\in I} \frac{\varphi(i)^p}{\psi(i)^p}\, di\right)^{-1/p} \end{cases}$$

*If* it worked, it would embody the enriched generalized logic as a functor

$$\mathbf{LT}^p : \mathbf{Prob}^{\mathrm{op}} \longrightarrow [0,\infty]_\otimes\text{-}\mathbf{Cat}$$

and quantifiers would be given as adjoints to reindexing, justifying their definition.

# The mirage of an enriched hyperdoctrine

Fix $\pi_I : I \times J \to I$, an enriched left adjoint $\int_i^p \dashv \pi_I^*$ would satisfy

$$\text{for all } \varphi \in \mathbf{LT}^p(I \times J), \psi \in \mathbf{LT}^p(I), \quad \int_i^p \varphi \vdash_I \psi = \varphi \vdash_{I \times J} \pi_I^* \psi$$

# The mirage of an enriched hyperdoctrine

Fix $\pi_I : I \times J \to I$, an enriched left adjoint $\int_i^p \dashv \pi_I^*$ would satisfy

$$\text{for all } \varphi \in \mathbf{LT}^p(I \times J), \psi \in \mathbf{LT}^p(I), \quad \int_i^p \varphi \vdash_I \psi = \varphi \vdash_{I \times J} \pi_I^* \psi$$

which unpacks to

$$\int_{i \in I}^{-p} \left( \int_{j \in I}^{p} \varphi(i,j) \right) \multimap \psi(i) = \int_{\substack{i \in I \\ j \in J}}^{-p} \varphi(i,j) \multimap \psi(i)$$

which is true by the fundamental relation of harmonic sum.

Similarly, we would have $\pi_I^* \dashv \int^{-p}$.

# The mirage of an enriched hyperdoctrine

Fix $\pi_I : I \times J \to I$, an enriched left adjoint $\int_i^p \dashv \pi_I^*$ would satisfy

$$\text{for all } \varphi \in \mathbf{LT}^p(I \times J), \psi \in \mathbf{LT}^p(I), \quad \int_i^p \varphi \vdash_I \psi = \varphi \vdash_{I \times J} \pi_I^* \psi$$

which unpacks to

$$\int_{i \in I}^{-p} \left( \int_{j \in I}^{p} \varphi(i,j) \right) \multimap \psi(i) = \int_{\substack{i \in I \\ j \in J}}^{-p} \varphi(i,j) \multimap \psi(i)$$

which is true by the fundamental relation of harmonic sum.

Similarly, we would have $\pi_I^* \dashv \int^{-p}$.

**Notice: it's absolutely crucial that $\vdash_I$ is enriched!**

# The mirage of an enriched hyperdoctrine

Why doesn't it *just* work?

# The mirage of an enriched hyperdoctrine

Why doesn't it *just* work?

**Entailment isn't transitive.** It's worth to see in detail why (for $p = 1$ for simplicity):

$$(\varphi \vdash_I \psi) \otimes (\psi \vdash_I \sigma) \leq (\varphi \vdash_I \sigma) \iff (\varphi \vdash_I \psi)^{-1} \otimes (\psi \vdash_I \sigma)^{-1} \geq (\varphi \vdash_I \sigma)^{-1}$$

Thus

$$\int_{i \in I} \int_{i' \in I} \frac{\varphi(i)}{\psi(i)} \frac{\psi(i')}{\sigma(i')}\, di\, di' \geq \int_{i \in \Delta_I} \frac{\varphi(i)}{\cancel{\psi(i)}} \frac{\cancel{\psi(i)}}{\sigma(i)}\, {\color{red} di\, di} \not\geq \int_{i \in I} \frac{\varphi(i)}{\sigma(i)}\, {\color{red} di}$$

# The mirage of an enriched hyperdoctrine

Why doesn't it *just* work?

**Entailment isn't transitive.** It's worth to see in detail why (for $p = 1$ for simplicity):

$$(\varphi \vdash_I \psi) \otimes (\psi \vdash_I \sigma) \leq (\varphi \vdash_I \sigma) \iff (\varphi \vdash_I \psi)^{-1} \otimes (\psi \vdash_I \sigma)^{-1} \geq (\varphi \vdash_I \sigma)^{-1}$$

Thus

$$\int_{i \in I} \int_{i' \in I} \frac{\varphi(i)}{\psi(i)} \frac{\psi(i')}{\sigma(i')}\, di\, di' \geq \int_{i \in \Delta_I} \frac{\varphi(i)}{\cancel{\psi(i)}} \frac{\cancel{\psi(i)}}{\sigma(i)}\, \color{red}{di\, di} \not\geq \int_{i \in I} \frac{\varphi(i)}{\sigma(i)}\, \color{red}{di}$$

So there seems to be a metalevel issue in the way we combine these numbers: if $\Delta_I$ was a subset of $\left(I \times I, \frac{di di'}{\sqrt{di di'}}\right)$, there would be no problem!

# The mirage of an enriched equipment of relations

An equipment of relations is an alternative way to present a first-order logic.

**Definition**. Define $[0, \infty]_{\otimes}\mathbf{Rel}$ as the $[0, \infty]_{\otimes}$-enriched equipment with prob. spaces & their maps as objects & tight maps, enriched relations as loose arrows, and squares given by

$$\begin{array}{ccc} A & \overset{R}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;} & B \\ f\downarrow & & \downarrow g \\ C & \underset{S}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;} & D \end{array} \quad := \quad R \vdash_{A \times B} S(f, g) \quad \in \quad [0, \infty]_{\otimes}.$$

# The mirage of an enriched equipment of relations

An equipment of relations is an alternative way to present a first-order logic.

**Definition**. Define $[0,\infty]_{\otimes}\mathbf{Rel}$ as the $[0,\infty]_{\otimes}$-enriched equipment with prob. spaces & their maps as objects & tight maps, enriched relations as loose arrows, and squares given by

$$\begin{array}{ccc} A & \overset{R}{\longrightarrow\!\!\!\!\!\!\!\!\!\mapstochar\;\;\;\;} & B \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ C & \underset{S}{\longrightarrow\!\!\!\!\!\!\!\!\!\mapstochar\;\;\;\;} & D \end{array} \quad := \quad R \vdash_{A\times B} S(f,g) \qquad \in \ [0,\infty]_{\otimes}.$$

There being '$a$' square means $1 \leq \begin{array}{ccc} A & \overset{R}{\nrightarrow} & B \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ C & \underset{S}{\nrightarrow} & D \end{array}$, in which case we draw an arrow inside.

Note this is parameterized by a choice of filter, but we fixed $F = [1,\infty]$.

# The mirage of an enriched equipment of relations

The existential quantifier is involved in composing relations:

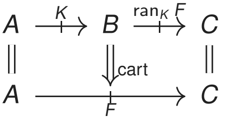
$$A \overset{R}{\nrightarrow} B \overset{S}{\nrightarrow} C, \quad (R \odot S)(a, c) := \int_{b \in B}^{p} R(a, b) \otimes S(b, c).$$

# The mirage of an enriched equipment of relations

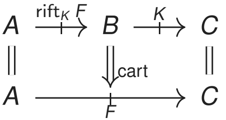
The existential quantifier is involved in composing relations:

$$A \overset{R}{\nrightarrow} B \overset{S}{\nrightarrow} C, \quad (R \odot S)(a,c) := \int_{b \in B}^{p} R(a,b) \otimes S(b,c).$$

The universal quantifier gives right lifts and extensions:




$$\mathrm{ran}_K F(b,c) = \int_{a}^{-p} K(a,b) \multimap F(a,c),$$




$$\mathrm{rift}_K F(a,b) = \int_{c \in C}^{-p} K(b,c) \multimap F(a,c).$$

# The mirage of an enriched equipment of relations

There are also all restrictions, hence companions and conjoints:

$$\begin{array}{ccc} A & \overset{S(f,g)}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;\;} & B \\ {\scriptstyle f}\big\downarrow & \Downarrow \text{cart} & \big\downarrow{\scriptstyle g} \\ C & \underset{S}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;\;} & D \end{array}, \qquad S(f,g)(a,b) := S(f(a), g(b)),$$

**Example**. Take $[0,1]$ with its uniform distribution, for $f : X \to [0,1]$ as before we have:

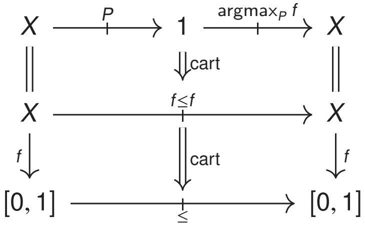
$$\begin{array}{ccc} X & \overset{f \multimap f}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;\;} & X \\ {\scriptstyle f}\big\downarrow & \Downarrow \text{cart} & \big\downarrow{\scriptstyle f} \\ [0,1] & \underset{\multimap}{\longrightarrow\!\!\!\!\!\!\!\!+\;\;\;\;} & [0,1] \end{array}$$

Incidentally, $(f \multimap f)(x, x') = 1 \wedge f(x)/f(x')$ appears as in the Metropolis–Hastings' Monte-Carlo sampling method.

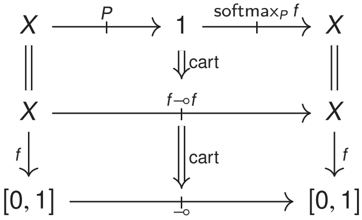
# The mirage of a perfect correspondance of argmax and softmax

Argmax and softmax are the same operation performed in different worlds:

In the $\mathbf{Prop}_\wedge$-equipment **Rel**:




In the $[0, \infty]_\otimes$-equipment $[0, \infty]_\otimes$**Rel**:




Here we are exhibiting *constrained argmax*, i.e. argmax subject to an optimization constraint $P$ given by a predicate. Put $P = 1$ for vanilla argmax.

# Conclusions

# Takeaways

1. $[0, \infty]_{\otimes}$ are a very rich setting!

# Takeaways

1. $[0, \infty]_\otimes$ are a very rich setting!
2. It's crucial to work enriched to get enriched results

# Takeaways

1. $[0, \infty]_{\otimes}$ are a very rich setting!
2. It's crucial to work enriched to get enriched results
3. **Universal properties reflect meta-level structure**, and thus the more structure in the enrichment base, the wilder the zoo of internally-reflected structure

# Takeaways

1. $[0, \infty]_\otimes$ are a very rich setting!
2. It's crucial to work enriched to get enriched results
3. **Universal properties reflect meta-level structure**, and thus the more structure in the enrichment base, the wilder the zoo of internally-reflected structure
4. $[0, \infty]_\otimes$-enriched universal properties are *equations*, and one can compute universal objects by solving them

# Takeaways

1. $[0, \infty]_{\otimes}$ are a very rich setting!
2. It's crucial to work enriched to get enriched results
3. **Universal properties reflect meta-level structure**, and thus the more structure in the enrichment base, the wilder the zoo of internally-reflected structure
4. $[0, \infty]_{\otimes}$-enriched universal properties are *equations*, and one can compute universal objects by solving them
5. Lack of idempotency keeps biting back

# Open questions

1. Is ‘very linear’ logic a thing? What’s its relationship to $\mathbb{PL}$?

# Open questions

1. Is ‘very linear’ logic a thing? What’s its relationship to $\mathbb{PL}$?
2. How do we fix the enriched Lindenbaum–Tarski algebras?

   Owen Lynch and David Jaz Myers’ have intriguing ideas about more sophisticated forms of enrichment which might save the day.

# Open questions

1. Is ‘very linear’ logic a thing? What’s its relationship to $\mathbb{PL}$?
2. How do we fix the enriched Lindenbaum–Tarski algebras?
   Owen Lynch and David Jaz Myers’ have intriguing ideas about more sophisticated forms of enrichment which might save the day.
3. Why are probability spaces natural domains of quantification?

# Open questions

1. Is 'very linear' logic a thing? What's its relationship to $\mathbb{PL}$?
2. How do we fix the enriched Lindenbaum–Tarski algebras?

   Owen Lynch and David Jaz Myers' have intriguing ideas about more sophisticated forms of enrichment which might save the day.
3. Why are probability spaces natural domains of quantification?
4. Can we motivate other constructions in analysis and probability theory from a logical/categorical POV?

   e.g. $L^p$ spaces, mutual information, Giry monad, etc.

# Thanks for your attention!

# References I


G. Bacci, R. Mardare, P. Panangaden, and G. Plotkin, *Propositional Logics for the Lawvere Quantale*, arXiv:2302.01224 [cs], Feb. 2023. DOI: 10.48550/arXiv.2302.01224. [Online]. Available: http://arxiv.org/abs/2302.01224 (visited on 03/01/2023).

——, *Polynomial Lawvere Logic*, en, arXiv:2402.03543 [cs], Feb. 2024. [Online]. Available: http://arxiv.org/abs/2402.03543 (visited on 02/12/2024).

M. Barr, **-Autonomous Categories*, ser. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 1979, vol. 752, ISBN: 978-3-540-09563-7 978-3-540-34850-4. DOI: 10.1007/BFb0064579. [Online]. Available: http://link.springer.com/10.1007/BFb0064579 (visited on 03/20/2024).

# References II


N. Blanco and N. Zeilberger, “Bifibrations of Polycategories and Classical Linear Logic”, *Electronic Notes in Theoretical Computer Science*, The 36th Mathematical Foundations of Programming Semantics Conference, 2020, vol. 352, pp. 29–52, Oct. 2020, ISSN: 1571-0661. DOI: `10.1016/j.entcs.2020.09.003`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1571066120300499` (visited on 03/20/2024).

R. Cockett and R. Seely, “Weakly distributive categories”, *Journal of Pure and Applied Algebra*, vol. 114, pp. 133–173, Nov. 1997. DOI: `10.1016/0022-4049(95)00160-3`.

F. Dagnino and F. Pasquali, *Quantitative Equality in Substructural Logic via Lipschitz Doctrines*, arXiv:2110.05388 [cs, math], Nov. 2022. DOI: `10.48550/arXiv.2110.05388`. [Online]. Available: `http://arxiv.org/abs/2110.05388` (visited on 03/01/2023).

# References III


F. W. Lawvere, “Adjointness in foundations”, *Dialectica*, pp. 281–296, 1969, Publisher: JSTOR. [Online]. Available: https://www.jstor.org/stable/42969800?casa_token=GL45s_dCKqQAAAAA:SV1lSO9MTQgPgPiwpYRwbX2Dt8agvDzsifxWybePlKcVBm7MqgcDrfWueuE94LKuuKEutiHxK1tb I4c4KNg760knXF (visited on 03/22/2024).

——,“Metric spaces, generalized logic, and closed categories”, *Rendiconti del seminario matématico e fisico di Milano*, vol. 43, pp. 135–166, 1973, Publisher: Springer.

F. Loregian, *(Co)end Calculus*, ser. London Mathematical Society Lecture Note Series. Cambridge: Cambridge University Press, 2021, ISBN: 978-1-108-74612-0. DOI: 10.1017/9781108778657. [Online]. Available: https://www.cambridge.org/core/books/coend-calculus/C662E90767358B336F17B606D19D8C43 (visited on 02/24/2023).

# References IV


- D. S. Mitrinovic and P. M. Vasic, *Analytic inequalities*. Springer, 1970, vol. 61.
- P. Perrone and W. Tholen, *Kan extensions are partial colimits*, en, arXiv:2101.04531 [math], Feb. 2021. [Online]. Available: http://arxiv.org/abs/2101.04531 (visited on 03/22/2024).
- M. Shulman, “LNL polycategories and doctrines of linear logic”, en, *Logical Methods in Computer Science*, vol. Volume 19, Issue 2, p. 7662, Apr. 2023, ISSN: 1860-5974. DOI: 10.46298/lmcs-19(2:1)2023. [Online]. Available: https://lmcs.episciences.org/7662 (visited on 03/19/2024).
- R. Street, “Frobenius monads and pseudomonoids”, *Journal of Mathematical Physics*, vol. 45, no. 10, pp. 3930–3948, Oct. 2004, ISSN: 0022-2488. DOI: 10.1063/1.1788852. [Online]. Available: https://doi.org/10.1063/1.1788852 (visited on 03/20/2024).