Towards Logical Foundations of Neuro-symbolic Methods

Matteo Capucci

University of Strathclyde / Independent Researcher (ARIA Creator)

August 03rd, 2026 — CAIA seminar

Acknowledgments

Matteo
Capucci

Bob
Atkey

Katya
Komendantskaya

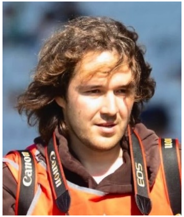
Charles
Grellois

Matthew
Daggitt

Adequate Losses via Quantitative Linear Logic

(`http://arxiv.org/abs/2605.13348`)

Acknowledgments



Thomas Flinkow

Katya Komendantskaya

Matteo Capucci

Rosemary Monahan

Quantitative Linear Logic for Neuro-Symbolic Learning and Verification

(https://arxiv.org/abs/2605.13845)

Property-driven training

Goal: train a neural model $f_\theta : X \longrightarrow \mathbb{R}^n$ to satisfy a specification φ

Property-driven training

Goal: train a neural model $f_\theta : X \longrightarrow \mathbb{R}^n$ to satisfy a specification φ

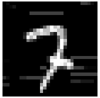
e.g.

Property-driven training

Goal: train a neural model $f_\theta : X \longrightarrow \mathbb{R}^n$ to satisfy a specification φ

e.g.

- **classification robustness**



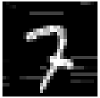
$$\forall x \in X,\ |x - \hat{x}_i| < \delta \implies |f_\theta(x) - f_\theta(\hat{x}_i)| < \epsilon$$

Property-driven training

Goal: train a neural model $f_\theta : X \longrightarrow \mathbb{R}^n$ to satisfy a specification φ

e.g.

- **classification robustness**



$$\forall x \in X,\ |x - \hat{x}_i| < \delta \implies |f_\theta(x) - f_\theta(\hat{x}_i)| < \epsilon$$

- **inductive biases**

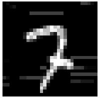
$$\forall x \in X,\ |x - \hat{x}_i| < \delta \implies \bigwedge_{(i,j) \in P} (\neg f_\theta(x)_i \vee \neg f_\theta(x)_j) \wedge (f_\theta(x)_i \vee f_\theta(x)_j)$$

Property-driven training

Goal: train a neural model $f_\theta : X \rightarrow \mathbb{R}^n$ to satisfy a specification φ

e.g.

- **classification robustness**

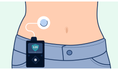


$$\forall x \in X,\ |x - \hat{x}_i| < \delta \implies |f_\theta(x) - f_\theta(\hat{x}_i)| < \epsilon$$

- **inductive biases**

$$\forall x \in X,\ |x - \hat{x}_i| < \delta \implies \bigwedge_{(i,j)\in P} (\neg f_\theta(x)_i \vee \neg f_\theta(x)_j) \wedge (f_\theta(x)_i \vee f_\theta(x)_j)$$

- **safety constraints**



$$\forall x \in X,\ C(t) + C\left(f_\theta(x), \frac{\ln k_a - \ln k_e}{k_a - k_e}\right) \leq C_{\text{safe}},$$ (Daggitt et al., 2026)

Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \to (\Theta \to \mathbb{R})$$

$$\mathrm{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

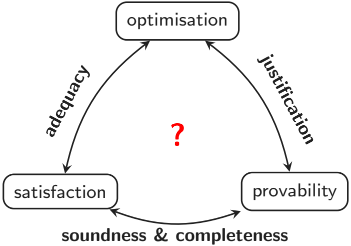
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \to (\Theta \to \mathbb{R})$$

$$\mathrm{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



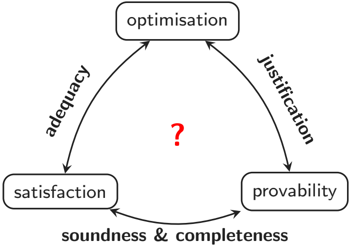
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \to (\Theta \to \mathbb{R})$$

$$\mathrm{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



- logical properties,

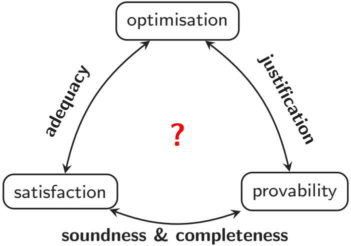
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \longrightarrow (\Theta \longrightarrow \mathbb{R})$$

$$\mathrm{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



- logical properties,
- analytical properties

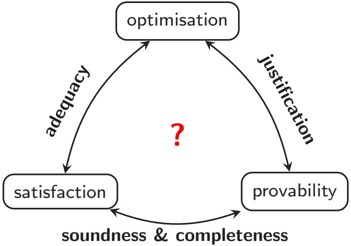
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \longrightarrow (\Theta \longrightarrow \mathbb{R})$$

$$\mathrm{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



- logical properties,
- analytical properties
- adequacy: $[\![\varphi]\!] \rightsquigarrow f_\theta \models \varphi$?

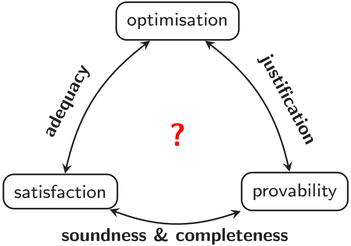
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \to (\Theta \to \mathbb{R})$$

$$\mathrm{argmin}_\theta \; \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\le x\le b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



effectiveness

- logical properties,
- analytical properties
- adequacy: $[\![\varphi]\!] \leftrightsquigarrow f_\theta \models \varphi$?

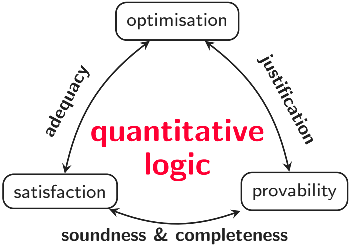
Property-driven training

Strategy: encode φ in the training objective via a Differentiable Logic (DL)

$$[\![-]\!] : \mathrm{Frm} \longrightarrow (\Theta \longrightarrow \mathbb{R})$$

$$\operatorname{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}),\hat{y}) + \lambda \sup_{a\leq x\leq b} [\![\varphi]\!](f_\theta(x),\hat{y})\right]$$

A challenging balancing act...



- logical properties,

effectiveness $\left\{\right.$
- analytical properties
- adequacy: $[\![\varphi]\!] \leftrightsquigarrow f_\theta \models \varphi$?

What is
Quantitative Logic?

What is Quantitative Logic?

Following **Lawvere**,

quantitative logic has **quantitative** syntactic / semantic validity

[1] but also: **continuous model theory** (Yaacov, Berenstein, etc.), or the **logics of the Lawvere quantale** (Bacci, Mardare, Panangaden, Plotkin), etc.

What is Quantitative Logic?

Following **Lawvere**,

$$\textbf{quantitative}\text{ logic has }\textbf{quantitative}\ \underbrace{\text{syntactic}}_{\textbf{provability}} \,/\, \underbrace{\text{semantic}}_{\textbf{satisfaction}}\ \text{validity}$$

[1]but also: **continuous model theory** (Yaacov, Berenstein, etc.), or the **logics of the Lawvere quantale** (Bacci, Mardare, Panangaden, Plotkin), etc.

What is Quantitative Logic?

Following **Lawvere**,

quantitative logic has **quantitative** $\underbrace{\text{syntactic}}_{\textbf{provability}}$ / $\underbrace{\text{semantic}}_{\textbf{satisfaction}}$ validity

$\neq$ *logics* **about** *quantity*, such as **fuzzy logics**[1], in which only the semantics is quantitative.

[1] but also: **continuous model theory** (Yaacov, Berenstein, etc.), or the **logics of the Lawvere quantale** (Bacci, Mardare, Panangaden, Plotkin), etc.

What is Quantitative Logic?

Following **Lawvere**,

quantitative logic has **quantitative** $\underbrace{\text{syntactic}}_{\textbf{provability}}$ / $\underbrace{\text{semantic}}_{\textbf{satisfaction}}$ validity

$\neq$ *logics* **about** *quantity*, such as **fuzzy logics**[1], in which only the semantics is quantitative.

Advantages:

- can axiomatize soft connectives,
- can makes sense of adequacy in property-driven training

[1] but also: **continuous model theory** (Yaacov, Berenstein, etc.), or the **logics of the Lawvere quantale** (Bacci, Mardare, Panangaden, Plotkin), etc.

The big picture

Bayesian probability, information theory,
machine learning, statistics, etc.

To give a logical treatmeant of quantitative disciplines
which have only been treated as logical informally, if at all.

The big picture

Bayesian probability, information theory,
machine learning, statistics, etc.

To give a logical treatmeant of quantitative disciplines which have only been treated as logical informally, if at all.

In particular: give logical foundations to neuro-symbolic methods by giving logical meaning to the loss

The multiplicative reals

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

	polarity	additive	multiplicative
duality	positive		
	negative		

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

	polarity	additive	multiplicative
duality	**positive**		$\mathbf{1} := 1$ $a \otimes b$
	negative		

where

$\otimes$ is multiplication,

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

distributes over

	polarity	**additive**	**multiplicative**
duality	**positive**		$\mathbf{1} := 1$ $a \otimes b$
	negative		

where

$\otimes$ is multiplication,

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

distributes over

	polarity	**additive**	**multiplicative**
duality	**positive**	$\bot := 0$ $a \oplus b$	$\mathbf{1} := 1$ $a \otimes b$
	negative		

where

$\otimes$ is multiplication,

$\oplus$ is addition,

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

distributes over

	polarity	**additive**	**multiplicative**
duality $a^* := 1/a$	**positive**	$\bot := 0$ $a \oplus b$	$\mathbf{1} := 1$ $a \otimes b$
	negative		

where

$\otimes$ is multiplication,

$\oplus$ is addition,

The multiplicative reals

$[0,\infty]_{\oplus,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

distributes over

	polarity	additive	multiplicative
duality $a^* := 1/a$ de Morgan	**positive**	$\perp := 0$ $a \oplus b$	$\mathbf{1} := 1$ $a \otimes b$ $\infty \otimes 0 := 0$
	negative	$\top := \infty$ $a \oplus^* b$	$\mathbf{1} := 1$ $a \otimes^* b$ $\infty \otimes^* 0 := \infty$

where

- $\otimes$ and $\otimes^*$ are multiplication (conjunctive and disjunctive),
- $\oplus$ is addition,
- $\oplus^*$ is harmonic sum: $a \oplus^* b = (a^* \oplus b^*)^*$,

The multiplicative p-reals

$[0,\infty]_{\oplus^p,\otimes}$ with the following algebraic structure on $([0,\infty],\leq)$:

distributes over

	polarity	additive			multiplicative	
duality $a^* := 1/a$ (de Morgan)	**positive**	$\bot := 0$, $a \oplus^p b$	$\xrightarrow{p\to+\infty}$	$\bot := 0$, $a \vee b$	$\mathbf{1} := 1$, $a \otimes b$	$\infty \otimes 0 := 0$
	negative	$\top := \infty$, $a \oplus^{-p} b$	$\xrightarrow{p\to+\infty}$	$\top := \infty$, $a \wedge b$	$\mathbf{1} := 1$, $a \otimes^* b$	$\infty \otimes^* 0 := \infty$
		soft		hard		

where

$\otimes$ and $\otimes^*$ are multiplication (conjunctive and disjunctive),

$\oplus^p$ is p-sum: $a \oplus^p b = (a^p \oplus b^p)^{1/p}$, $0 < p \leq \infty$,

$\oplus^{-p}$ is harmonic p-sum: $a \oplus^{-p} b = (a^* \oplus^p b^*)^*$, QLL

pQLL is a quantitative logic in which

all notions of (proof, sequent) validity are valued in $[0, \infty]_{\oplus^p, \otimes}$.

pQLL is a quantitative logic in which
all notions of (proof, sequent) validity are valued in $[0,\infty]_{\oplus^p,\otimes}$.

This means

1. the **proof calculus** is quantitative and
2. the **model theory** is quantitative.

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{A_1, \ldots, A_n}_{\text{given \textbf{all} of these...}} \quad \vdash \quad \underbrace{B_1, \ldots B_m}_{\text{...\textbf{one} of these holds}}$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given } \textbf{all} \text{ of these...}} \vdash \underbrace{\Delta}_{\text{...}\textbf{one} \text{ of these holds}}$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given \textbf{all} of these...}} \vdash \underbrace{\Delta}_{\text{...\textbf{one} of these holds}}$$

Rules combine the premises to get the conclusion

$$\frac{\Gamma \vdash A \qquad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given \textbf{all} of these...}} \vdash \underbrace{\Delta}_{\text{...\textbf{one} of these holds}}$$

Rules combine the premises to get the conclusion

$$\frac{\Gamma \vdash A \qquad \text{and} \qquad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given \textbf{all} of these...}} \vdash \underbrace{\Delta}_{\text{...\textbf{one} of these holds}}$$

Rules combine the premises to get the conclusion, **axioms** introduce sequents:

$$\frac{\mathrm{T}}{A \vdash A} \quad , \qquad \frac{\Gamma \vdash A \quad \text{and} \quad A \vdash \Delta}{\Gamma \vdash \Delta} \quad , \qquad \ldots$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given \textbf{all} of these...}} \vdash \underbrace{\Delta}_{\text{...\textbf{one} of these holds}}$$

Rules combine the premises to get the conclusion, **axioms** introduce sequents:

$$\frac{\color{red}{\mathrm{T}}}{A \vdash A}\ , \qquad \frac{\Gamma \vdash A \quad \color{red}{\text{and}} \quad A \vdash \Delta}{\Gamma \vdash \Delta}\ , \qquad \ldots$$

Proofs are trees: they are valid is they can be written with the given rules

$$\frac{\dfrac{\color{red}{\mathrm{T}}}{A \vdash A} \quad \color{red}{\text{and}} \quad \dfrac{\color{red}{\mathrm{T}}}{A \vdash A}}{A \vee A \vdash A}$$

Sequent Calculus 101

A **sequent** is just a pair of lists of formulae:

$$\underbrace{\Gamma}_{\text{given } \textbf{all} \text{ of these...}} \quad \vdash \quad \underbrace{\Delta}_{\text{...} \textbf{one} \text{ of these holds}}$$

Rules combine the premises to get the conclusion, **axioms** introduce sequents:

$$\frac{\top}{A \vdash A} \quad , \qquad \frac{\Gamma \vdash A \quad \text{and} \quad A \vdash \Delta}{\Gamma \vdash \Delta} \quad , \qquad \ldots$$

Proofs are trees: they are valid is they can be written with the given rules

$$\frac{\dfrac{\top}{A \vdash A} \quad \text{and} \quad \dfrac{\top}{A \vdash A}}{A \vee A \vdash A}$$

A sequent is **provable** if there exists a proof ending with it.

Quantitative Sequent Calculus 101

Rules must explicitly say **how** to combine the premises to get the conclusion, axioms must come with a validity $\in [0, \infty]_{\oplus,\otimes}$, e.g.

$$\frac{1}{A \vdash A} \qquad\qquad \frac{\Gamma \vdash A \quad \otimes \quad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Quantitative Sequent Calculus 101

Rules must explicitly say **how** to combine the premises to get the conclusion, axioms must come with a validity $\in [0,\infty]_{\oplus,\otimes}$, e.g.

$$\frac{1}{A \vdash A} \qquad\qquad \frac{\Gamma \vdash A \quad \otimes \quad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Then proofs get a validity obtained by traversing the tree:

$$[0,\infty] \ni \left| \begin{array}{c} \vdots\, \pi \\ \Gamma \vdash \Delta \end{array} \right|$$

Quantitative Sequent Calculus 101

Rules must explicitly say **how** to combine the premises to get the conclusion, axioms must come with a validity $\in [0, \infty]_{\oplus,\otimes}$, e.g.

$$\frac{1}{A \vdash A} \qquad\qquad \frac{\Gamma \vdash A \quad \otimes \quad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Then proofs get a validity obtained by traversing the tree:

$$[0, \infty] \ni \left| \begin{array}{c} \vdots\, \pi \\ \Gamma \vdash \Delta \end{array} \right| = \begin{cases} r & \text{if} \quad \pi = \dfrac{r}{\Gamma \vdash \Delta} \\ & \\ & \end{cases}$$

Quantitative Sequent Calculus 101

Rules must explicitly say **how** to combine the premises to get the conclusion, axioms must come with a validity $\in [0,\infty]_{\oplus,\otimes}$, e.g.

$$\frac{1}{A \vdash A} \qquad\qquad \frac{\Gamma \vdash A \quad \otimes \quad A \vdash \Delta}{\Gamma \vdash \Delta}$$

Then proofs get a validity obtained by traversing the tree:

$$[0,\infty] \ni \left| \begin{array}{c} \vdots\ \pi \\ \Gamma \vdash \Delta \end{array} \right| = \begin{cases} r & \text{if} \quad \pi = \dfrac{r}{\Gamma \vdash \Delta} \\[2em] |\pi_1| * |\pi_2| & \text{if} \quad \pi = \dfrac{\begin{array}{c} \vdots\ \pi_1 \\ \Gamma_1 \vdash \Delta_1 \end{array} \quad * \quad \begin{array}{c} \vdots\ \pi_2 \\ \Gamma_1 \vdash \Delta_1 \end{array}}{\Gamma \vdash \Delta} \end{cases}$$

Quantitative sequent calculi

The **provability** of a sequent is thus defined in terms of validities of its proofs:

$$|\Gamma \vdash \Delta| = \bigvee \left\{ |\pi| \;\middle|\; \begin{array}{c} \vdots\, \pi \\ \Gamma \vdash \Delta \end{array} \right\}$$

Note that when $\mathcal{V} = \mathbf{Prop}$, this recovers the classical definition.

pQLL

By quantitifying **linear logic** (MALL) we obtain pQLL:

Structural Rules

$$\frac{1}{A \vdash A}\ \mathsf{AX} \qquad \frac{\Gamma \vdash A \quad \otimes \quad A \vdash \Delta}{\Gamma \vdash \Delta}\ \mathsf{CUT}$$

$$\frac{1}{\vdash}\ \mathsf{EMP} \qquad \frac{\Gamma \vdash \Delta \quad \otimes \quad \Gamma' \vdash \Delta'}{\Gamma, \Gamma' \vdash \Delta, \Delta'}\ \mathsf{MIX}$$

$$\frac{0}{\Gamma \vdash \Delta}\ \mathsf{EFQ}$$

Multiplicative Connectives

$$\frac{1}{\mathbf{1} \vdash}\ \mathbf{1}_{\mathsf{L}}$$

$$\frac{A \vdash \Delta \quad \otimes \quad B \vdash \Delta'}{A \parr B \vdash \Delta, \Delta'}\ \parr_{\mathsf{L}}$$

$$\frac{A, B \vdash \Delta}{A \otimes B \vdash \Delta}\ \otimes_{\mathsf{L}}$$

Additive Connectives

$$\frac{\infty}{\Gamma, \bot \vdash \Delta}\ \bot_{\mathsf{L}}$$

$$\frac{A \vdash \Delta \quad \oplus^{p} \quad B \vdash \Delta}{A \wedge B \vdash \Delta}\ \wedge_{\mathsf{L}}$$

$$\frac{A \vdash \Delta \quad \oplus^{-p} \quad B \vdash \Delta}{A \vee B \vdash \Delta}\ \vee_{\mathsf{L}}$$

Duality

$$\frac{\Gamma \vdash A, \Delta}{\Gamma, A^{\perp} \vdash \Delta}\ (-)^{\perp}_{\mathsf{L}} \qquad \frac{\Gamma, A \vdash \Delta}{\Gamma \vdash A^{\perp}, \Delta}\ (-)^{\perp}_{\mathsf{R}}$$

One proof

$$\dfrac{\dfrac{1}{A \vdash A}\,\mathsf{AX} \qquad \oplus^{-p} \qquad \dfrac{1}{A \vdash A}\,\mathsf{AX}}{A \vee A \vdash A}\,\vee_{\mathsf{L}}$$

One proof

$$\left|\dfrac{\dfrac{1}{A \vdash A}\,\mathsf{AX} \quad \oplus^{-p} \quad \dfrac{1}{A \vdash A}\,\mathsf{AX}}{A \vee A \vdash A}\,\vee_{\mathsf{L}}\right| = \left|\dfrac{1}{A \vdash A}\,\mathsf{AX}\right| \oplus^{-p} \left|\dfrac{1}{A \vdash A}\,\mathsf{AX}\right| = 1 \oplus^{-p} 1 = 2^{-1/p}$$

One proof

$$\left| \dfrac{\dfrac{1}{A \vdash A}\,\mathsf{AX} \quad \oplus^{-p} \quad \dfrac{1}{A \vdash A}\,\mathsf{AX}}{A \vee A \vdash A}\,\vee_{\mathsf{L}} \right| = \left| \dfrac{1}{A \vdash A}\,\mathsf{AX} \right| \oplus^{-p} \left| \dfrac{1}{A \vdash A}\,\mathsf{AX} \right| = 1 \oplus^{-p} 1 = 2^{-1/p}$$

Theorem. $\Gamma \vdash \Delta$ is provable in MALL if and only if $|\Gamma \vdash \Delta|_{p\mathsf{QLL}} > 0$.

Proof-theoretic results

CUT-elimination Theorem (Capucci, Atkey, Grellois, Komendantskaya, and Daggitt 2026)
CUT is eliminable, that is $|\Gamma \vdash \Delta|_{p\mathsf{QLL}} = |\Gamma \vdash \Delta|_{p\mathsf{QLL}^{\cancel{\mathsf{CUT}}}}$.

Proof sketch. The usual, paired with arguments that show that each proof rewrite is *validity non-decreasing*.

In the proof, we use that $\otimes$ distributes over $\oplus^p/\oplus^{-p}$, unitality and associativity of $\otimes$, etc.

Corollary. pQLL enjoys the subformula property, it is consistent ($|\vdash \bot| = 0$), and decidable.

Model theory

pQLL is **sound and complete** for a class of $[0, \infty]_{\otimes}$-enriched preorders called **softales**:

Definition. A **softale** is a set S equipped with an **enriched order** $\sqsubseteq : S \times S \to [0, \infty]$ such that

$$1 \leq (a \sqsubseteq a), \qquad (a \sqsubseteq b) \otimes (b \sqsubseteq c) \leq (a \sqsubseteq c)$$

together with operations $\bullet, \vee : S \times S \to S$, a duality $(-)^{\perp} : S^{\mathsf{op}} \to S$, satisfying certain algebraic laws...

Example: logspace

Call **logspace** the poset $([-\infty, +\infty], \geq)$ equipped with addition $(+)$ and (harmonic) **log-sum-exp**:

$$a \cup^1 b = -\log(e^{-a} + e^{-b}).$$

Example: logspace

Call **logspace** the poset $([-\infty, +\infty], \geq)$ equipped with addition $(+)$ and (harmonic) **log-sum-exp**:

$$a \cup^1 b = -\log(e^{-a} + e^{-b}).$$

	polarity	additive			multiplicative	
duality $a^* := -a$	**positive**	$\bot := +\infty$ $a \cap^p b$	$\xrightarrow[p\to\infty]{}$	$\bot := +\infty$ $a \cap^\infty b := \min(a, b)$	$\mathbf{1} := 0$ $a + b$	$-\infty + \infty := +\infty$
	negative	$\top := -\infty$ $a \cup^p b$	$\xrightarrow[p\to\infty]{}$	$\top := -\infty$ $a \cup^\infty b := \max(a, b)$	$\mathbf{1} := 0$ $a +^- b$	$-\infty +^- \infty := -\infty$

implication: $a \multimap b = a^* +^- b$, thus $b - a$, **enriched order**: $a \sqsubseteq b := e^{-(a \multimap b)}$,

Example: logspace

Call **logspace** the poset $([-\infty, +\infty], \geq)$ equipped with addition $(+)$ and (harmonic) **log-sum-exp**:

$$a \cup^1 b = -\log(e^{-a} + e^{-b}).$$

	polarity	additive	multiplicative
duality $a^* := -a$	**positive**	$\perp := +\infty$, $a \cap^p b$ $\xrightarrow[p\to\infty]{}$ $\perp := +\infty$, $a \cap^\infty b := \min(a, b)$	$\mathbf{1} := 0$, $a + b$, $-\infty + \infty := +\infty$
	negative	$\top := -\infty$, $a \cup^p b$ $\xrightarrow[p\to\infty]{}$ $\top := -\infty$, $a \cup^\infty b := \max(a, b)$	$\mathbf{1} := 0$, $a +^- b$, $-\infty +^- \infty := -\infty$

implication: $a \multimap b = a^* +^- b$, thus $b - a$, **enriched order**: $a \sqsubseteq b := e^{-(a \multimap b)}$,

This is the **natural model** for property-driven training: <u>losses live in logspace!</u>

Example: logspace

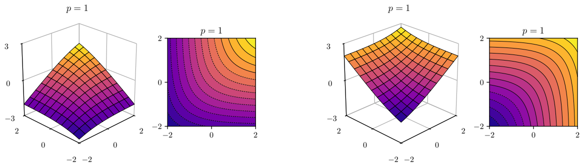
Call **logspace** the poset $([-\infty, +\infty], \geq)$ equipped with addition $(+)$ and (harmonic) **log-sum-exp**:

$$a \cup^1 b = -\log(e^{-a} + e^{-b}).$$

	polarity	**additive**			**multiplicative**	
duality $a^* := -a$	**positive**	$\perp := +\infty$ $a \cap^p b$	$\xrightarrow[p\to\infty]{}$	$\perp := +\infty$ $a \cap^\infty b := \min(a, b)$	$\mathbf{1} := 0$ $a + b$	$-\infty + \infty := +\infty$
	negative	$\top := -\infty$ $a \cup^p b$	$\xrightarrow[p\to\infty]{}$	$\top := -\infty$ $a \cup^\infty b := \max(a, b)$	$\mathbf{1} := 0$ $a +^- b$	$-\infty +^- \infty := -\infty$

implication: $a \multimap b = a^* +^- b$, thus $b - a$, **enriched order**: $a \sqsubseteq b := e^{-(a \multimap b)}$,

also, excellent gradients:



$$\frac{\partial(x \cup^1 y)}{\partial x} = \frac{e^y}{e^x + e^y},$$

$$\frac{\partial(x \cap^1 y)}{\partial x} = \frac{e^x}{e^x + e^y}$$

Application: property-driven verification

Consider $f_\theta : X \longrightarrow \mathbb{R}^n$ as in the beginning.

[2] plus MIX*, ask me later about it.

Application: property-driven verification

Consider $f_\theta : X \to \mathbb{R}^n$ as in the beginning.

Definition. The **theory of neuro-symbolic specifications** for f_θ at $(\hat{x}, \hat{y}) \in \mathcal{D}$ and $x \in X$ has set of propositional variables

$$K = \{\underbrace{\ulcorner \hat{y}_1 \urcorner, \ldots, \ulcorner \hat{y}_n \urcorner}_{\text{data labels}}, \underbrace{\ulcorner y_1 \urcorner, \ldots, \ulcorner y_n \urcorner}_{\text{model outputs}}\}$$

[2] plus MIX*, ask me later about it.

Application: property-driven verification

Consider $f_\theta : X \to \mathbb{R}^n$ as in the beginning.

Definition. The **theory of neuro-symbolic specifications** for f_θ at $(\hat{x}, \hat{y}) \in \mathcal{D}$ and $x \in X$ has set of propositional variables

$$K = \{\underbrace{\ulcorner \hat{y}_1 \urcorner, \ldots, \ulcorner \hat{y}_n \urcorner}_{\text{data labels}}, \underbrace{\ulcorner y_1 \urcorner, \ldots, \ulcorner y_n \urcorner}_{\text{model outputs}}\}$$

This amounts to the rules of pQLL[2] plus the following axioms

$$\frac{\hat{y}_i}{\vdash \ulcorner \hat{y}_i \urcorner} \qquad \frac{e^{-f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner} \qquad \frac{\hat{y}_i^*}{\vdash \ulcorner \hat{y}_i \urcorner^\perp} \qquad \frac{e^{f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner^\perp} \qquad \text{(grounding schema)}$$

[2]plus MIX*, ask me later about it.

Application: property-driven verification

Consider $f_\theta : X \to \mathbb{R}^n$ as in the beginning.

Definition. The **theory of neuro-symbolic specifications** for f_θ at $(\hat{x}, \hat{y}) \in \mathcal{D}$ and $x \in X$ has set of propositional variables

$$K = \{\underbrace{\ulcorner \hat{y}_1 \urcorner, \ldots, \ulcorner \hat{y}_n \urcorner}_{\text{data labels}}, \underbrace{\ulcorner y_1 \urcorner, \ldots, \ulcorner y_n \urcorner}_{\text{model outputs}}\}$$

This amounts to the rules of pQLL[2] plus the following axioms

$$\frac{\hat{y}_i}{\vdash \ulcorner \hat{y}_i \urcorner} \qquad \frac{e^{-f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner} \qquad \frac{\hat{y}_i^*}{\vdash \ulcorner \hat{y}_i \urcorner^\perp} \qquad \frac{e^{f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner^\perp} \qquad \text{(grounding schema)}$$

We denote such theory by $\mathbb{T}_{x,\hat{y},f_\theta}$.

[2] plus MIX*, ask me later about it.

Application: property-driven verification

Consider $f_\theta : X \longrightarrow \mathbb{R}^n$ as in the beginning.

Definition. The **theory of neuro-symbolic specifications** for f_θ at $(\hat{x}, \hat{y}) \in \mathcal{D}$ and $x \in X$ has set of propositional variables

$$K = \{\underbrace{\ulcorner \hat{y}_1 \urcorner, \ldots, \ulcorner \hat{y}_n \urcorner}_{\text{data labels}}, \underbrace{\ulcorner y_1 \urcorner, \ldots, \ulcorner y_n \urcorner}_{\text{model outputs}}\}$$

This amounts to the rules of pQLL[2] plus the following axioms

$$\frac{\hat{y}_i}{\vdash \ulcorner \hat{y}_i \urcorner} \qquad \frac{e^{-f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner} \qquad \frac{\hat{y}_i^*}{\vdash \ulcorner \hat{y}_i \urcorner^\perp} \qquad \frac{e^{f_\theta(x)_i}}{\vdash \ulcorner y_i \urcorner^\perp} \qquad \text{(grounding schema)}$$

We denote such theory by $\mathbb{T}_{x,\hat{y},f_\theta}$.

Example. Strong classification robustness: $\displaystyle\bigwedge_{i \in n} \hat{y}_i \multimap (y_i \multimap \delta)$

[2] plus MIX*, ask me later about it.

Definition. A **model** of $\mathbb{T}_{x,\hat{y},f_\theta}$ is a valuation of propositional variables in a softale:

$$[\![-]\!] : \mathbb{T}_{x,\hat{y},f_\theta} \longrightarrow S$$

Definition. A **model** of $\mathbb{T}_{x,\hat{y},f_\theta}$ is a valuation of propositional variables in a softale:

$$[\![-]\!] : \mathbb{T}_{x,\hat{y},f_\theta} \longrightarrow S$$

The **satisfaction** of a formula φ in the model $[\![-]\!]$ is thus

$$[\![-]\!] \models \varphi \;=\; \mathbf{1} \sqsubseteq_S [\![\varphi]\!] \quad \in [0,\infty]$$

Definition. A **model** of $\mathbb{T}_{x,\hat{y},f_\theta}$ is a valuation of propositional variables in a softale:

$$[\![-]\!] : \mathbb{T}_{x,\hat{y},f_\theta} \longrightarrow S$$

The **satisfaction** of a formula φ in the model $[\![-]\!]$ is thus

$$[\![-]\!] \models \varphi \;\; = \;\; \mathbf{1} \sqsubseteq_S [\![\varphi]\!] \quad \in [0, \infty]$$

Observation. There is a canonical model of $\mathbb{T}_{x,\hat{y},f_\theta}$ in logspace, defined by

$$[\![\ulcorner \hat{y}_i \urcorner]\!] = -\log \hat{y}_i, \qquad [\![\ulcorner y_i \urcorner]\!] = f_\theta(x)_i.$$

Soundness, completeness, & adequacy

Theorem. Given f_θ as above, given $(\hat{x}, \hat{y}) \in \mathcal{D}$ and for each $a \leq x \leq b$,

$$= \quad \hat{x}, \hat{y}, x, f_\theta \models \varphi \quad = \quad e^{-[\![\varphi]\!]}$$

Soundness, completeness, & adequacy

Theorem. Given f_θ as above, given $(\hat{x}, \hat{y}) \in \mathcal{D}$ and for each $a \leq x \leq b$,

$$= \underbrace{\hat{x}, \hat{y}, x, f_\theta \models \varphi \quad = \quad e^{-[\![\varphi]\!]}}_{\textbf{adequacy}}$$

$$\operatorname{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) + \lambda \sup_{a \leq x \leq b} [\![\varphi]\!](f_\theta(x), \hat{y})\right]$$

Soundness, completeness, & adequacy

Theorem. Given f_θ as above, given $(\hat{x}, \hat{y}) \in \mathcal{D}$ and for each $a \leq x \leq b$,

$$= \underbrace{\hat{x}, \hat{y}, x, f_\theta \models \varphi \quad = \quad e^{-[\![\varphi]\!]}}_{\textbf{adequacy}}$$

$$\operatorname{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\Big[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) + \lambda \sup_{a \leq x \leq b} -\log\big(\hat{x}, \hat{y}, x, f_\theta \models \varphi\big)\Big]$$

Soundness, completeness, & adequacy

Theorem. Given f_θ as above, given $(\hat{x}, \hat{y}) \in \mathcal{D}$ and for each $a \leq x \leq b$,

$$\overbrace{|\vdash_{\mathbb{T}_{x,\hat{y},f_\theta}} \varphi| \quad = \quad \hat{x}, \hat{y}, x, f_\theta \models \varphi}^{\textbf{soundness \& completeness}} \quad = \quad e^{-[\![\varphi]\!]}$$

$$\underbrace{\hat{x}, \hat{y}, x, f_\theta \models \varphi \quad = \quad e^{-[\![\varphi]\!]}}_{\textbf{adequacy}}$$

$$\operatorname{argmin}_\theta \ \mathbb{E}_{(\hat{x},\hat{y})\sim\mathcal{D}}\left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) + \lambda \sup_{a \leq x \leq b} -\log\left(\hat{x}, \hat{y}, x, f_\theta \models \varphi\right)\right]$$

Effectiveness

from (Flinkow, Komendantskaya, Capucci, and Monahan 2026):

Table 5: Selected experimental results for training and verifying models on MNIST, Fashion-MNIST, and Dice using different constraints and differentiable logics. Boldface indicates the best-performing configuration in each experiment, measured by product of PAcc and CSat. When a model is trained with a choice of ε, it is then attacked (CSec) and verified (CSat) for the same ε.

Constraint	Dataset	Logic	PAcc (%)	CSec$_\varepsilon$ (%)	CSat$_\varepsilon$ (%)	
					verified	unknown[a]
Clothing/Footwear (training with $\varepsilon = 0.1$)	Fashion-MNIST	Baseline	87.4 ± 0.5	35.8 ± 11.1[b]	19.0 ± 0.0	1.8 ± 2.5
		DL2	86.9 ± 0.8	99.8 ± 0.0	31.0 ± 19.0	36.6 ± 11.8
		$_5$STL	87.2 ± 0.5	99.8 ± 0.0	25.8 ± 11.9	35.0 ± 5.5
		$_5$QLL	**85.2 ± 1.7**	**99.0 ± 0.3**	**87.2 ± 12.3**	**12.0 ± 11.7**
Exactly-One (training with $\varepsilon = 4/255$)	Dice	Baseline	83.8 ± 4.7	25.6 ± 7.2[b]	6.2 ± 6.2	5.6 ± 3.0
		DL2	83.7 ± 2.1	90.0 ± 4.5	3.2 ± 3.0	6.5 ± 6.7
		Gödel	83.4 ± 1.6	85.8 ± 4.5	2.9 ± 4.2	10.3 ± 2.1
		Łukasiewicz	84.4 ± 3.2	92.2 ± 5.2	5.9 ± 6.2	8.1 ± 3.1
		Product	84.2 ± 1.9	92.2 ± 3.1	3.4 ± 3.4	10.3 ± 2.9
		$_2$Yager	85.9 ± 2.1	91.2 ± 6.7	3.4 ± 3.1	6.4 ± 0.8
		$_5$STL	82.9 ± 3.4	41.8 ± 12.7	10.6 ± 9.1	4.4 ± 2.1
		$_5$QLL	**82.5 ± 4.4**	**51.5 ± 14.4**	**14.7 ± 10.3**	**5.9 ± 5.4**
Strong classification robustness (training with $\varepsilon = 0.1$)	MNIST	Baseline	97.6 ± 0.4	19.3 ± 18.1[b]	2.0 ± 1.9	14.2 ± 18.1
		DL2	97.6 ± 0.4	100.0 ± 0.0	24.2 ± 10.3	29.2 ± 19.8
		$_5$STL	**97.5 ± 0.3**	**100.0 ± 0.0**	**93.4 ± 9.4**	**6.6 ± 9.4**
		$_5$QLL	**97.5 ± 0.3**	**100.0 ± 0.0**	**93.4 ± 9.4**	**6.6 ± 9.4**
Classification robustness (training with $\varepsilon = 0.2$)	MNIST	Baseline	97.6 ± 0.4	0.0 ± 0.0[b]	0.0 ± 0.0	0.0 ± 0.0
		DL2	97.1 ± 1.2	70.4 ± 34.0	10.4 ± 21.6	26.4 ± 32.8
		Gödel	97.5 ± 0.4	94.7 ± 1.6	0.0 ± 0.0	0.4 ± 0.5
		Łukasiewicz	97.6 ± 0.3	92.3 ± 5.1	0.0 ± 0.0	0.2 ± 0.4
		Product	97.6 ± 0.4	94.3 ± 3.0	0.0 ± 0.0	0.5 ± 0.6
		$_2$Yager	97.5 ± 0.6	93.3 ± 2.7	0.0 ± 0.0	0.2 ± 0.5
		$_5$STL	97.7 ± 0.5	68.5 ± 34.9	0.0 ± 0.0	4.0 ± 6.2
		$_5$QLL	**96.1 ± 0.8**	**67.7 ± 4.0**	**46.0 ± 23.2**	**30.0 ± 28.8**

[a] The ‘unknown’ column denotes verification instances that could neither be proved or disproved within a 30 s timeout.
[b] We tested baseline CSec against QLL, see Appendix A.6 for an explanation.

What's next

Logic

- First-order already in the workings! See where it all started: (Capucci 2024)

Logic

- First-order already in the workings! See where it all started: (Capucci 2024)
- Expand the logical understanding of the loss:

$$f_\theta \models \forall^0 (\hat{x}, \hat{y}) : \mathcal{D}.\ \left[\mathcal{L}_{\mathrm{pred}}(f_\theta(\hat{x}), \hat{y}) \ \wedge\ \lambda \forall^\infty x : (a, b).\ \varphi(f_\theta(x), \hat{y})\right]$$

What's next

Logic

- First-order already in the workings! See where it all started: (Capucci 2024)
- Expand the logical understanding of the loss:

$$f_\theta \models \forall^0 (\hat{x}, \hat{y}) : \mathcal{D}.\ \left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) \ \wedge \ \lambda \forall^\infty x : (a, b).\ \varphi(f_\theta(x), \hat{y})\right]$$

What's next

Logic

- First-order already in the workings! See where it all started: (Capucci 2024)
- Expand the logical understanding of the loss:

$$f_\theta \models \forall^0 (\hat{x}, \hat{y}) : \mathcal{D}.\ \left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) \ \wedge\ \lambda \forall^\infty x : (a, b).\ \varphi(f_\theta(x), \hat{y}) \right]$$

ML

- More thorough evaluation: larger models, larger specs

What's next

Logic

- First-order already in the workings! See where it all started: (Capucci 2024)
- Expand the logical understanding of the loss:

$$f_\theta \models \forall^0 (\hat{x}, \hat{y}) : \mathcal{D}.\ \left[\mathcal{L}_{\text{pred}}(f_\theta(\hat{x}), \hat{y}) \ \wedge \ \lambda \forall^\infty x : (a, b).\ \varphi(f_\theta(x), \hat{y}) \right]$$

ML

- More thorough evaluation: larger models, larger specs
- What does quantitative verification affords us?

Thanks!

Application: Bayesian probability

Every atomic probability space $p : \Omega \to [0,1]$ yields a theory with propositional constants $\ulcorner \omega \urcorner \in \Omega$ and axioms

$$\frac{p(a)}{\vdash a}\ a_{\mathsf{R}} \qquad \frac{p(a)^*}{a \vdash}\ a_{\mathsf{L}} \qquad \text{for all } a \in V.$$

Every event $A \subseteq \Omega$ corresponds to a formula in this theory, which we abbreviate as $\bigvee_{a \in A} a$, and denote again by A.

Application: Bayesian probability

Every atomic probability space $p : \Omega \to [0,1]$ yields a theory with propositional constants $\ulcorner \omega \urcorner \in \Omega$ and axioms

$$\frac{p(a)}{\vdash a}\ a_{\mathsf{R}} \qquad\qquad \frac{p(a)^*}{a \vdash}\ a_{\mathsf{L}} \qquad\qquad \text{for all } a \in V.$$

Every event $A \subseteq \Omega$ corresponds to a formula in this theory, which we abbreviate as $\bigvee_{a\in A} a$, and denote again by A. Then one can compute $p(B \mid A)$ by computing the odds $A \vdash A \cap B$:

$$\cfrac{\cfrac{\cfrac{\cfrac{\bigoplus_{b\in A\cap B} \bigoplus^*_{a\in A}\ p(a)^* \otimes p(b)}{\bigoplus_{b\in A\cap B} \bigoplus^*_{a\in A}\ (a \vdash \otimes \vdash b)}\ \bigoplus_b \bigoplus^*_a (a_{\mathsf{L}} \otimes b_{\mathsf{R}})}{\bigoplus_{b\in A\cap B} \bigoplus^*_{a\in A}\ a \vdash b}\ \bigoplus_b \bigoplus^*_a \mathsf{MIX}}{\bigoplus_{b\in A\cap B}\ A \vdash b}\ \bigoplus_b \vee_{\mathsf{L}}}{A \vdash A \cap B}\ \vee_{\mathsf{R}}$$

and we indeed have $\bigoplus_{b\in A\cap B} \bigoplus^*_{a\in A}\ p(a)^* \otimes p(b) = \dfrac{p(A \cap B)}{p(A)} = p(B \mid A)$.

References I

[1] M. Capucci, *On Quantifiers for Quantitative Reasoning*, Jun. 2024. DOI: 10.48550/arXiv.2406.04936. eprint: 2406.04936 (cs, math). Accessed: Jun. 26, 2024. [Online]. Available: http://arxiv.org/abs/2406.04936.

[2] M. Capucci, R. Atkey, C. Grellois, E. Komendantskaya, and M. Daggitt, "Adequate Losses via Quantitative Linear Logic", arXiv: 2605.13348 [cs.LO], Accessed: Jul. 18, 2026. [Online]. Available: http://arxiv.org/abs/2605.13348, pre-published.

[3] T. Flinkow, E. Komendantskaya, M. Capucci, and R. Monahan, "Quantitative Linear Logic for Neuro-Symbolic Learning and Verification", Accessed: May 14, 2026. [Online]. Available: https://arxiv.org/abs/2605.13845.

[4] F. W. Lawvere, "Metric spaces, generalized logic, and closed categories", *Rendiconti del seminario matématico e fisico di Milano*, vol. 43, pp. 135–166, 1973.