

Weighted Relational Models

Guy McCusker, University of Bath

DiALL, Marseille, May 2024

Weighted relational models

This talk will be an overview of *weighted relational models* of (differential) linear logic and typed lambda calculus.

Prehistory

- Girard's *quantitative semantics* introduced the idea of interpreting lambda-terms as power series.

Jean-Yves Girard, 'Normal Functors, Power Series and λ -Calculus', *Annals of Pure and Applied Logic* 37, no. 2 (1 February 1988): 129–77, [https://doi.org/10.1016/0168-0072\(88\)90025-5](https://doi.org/10.1016/0168-0072(88)90025-5).

- The relational model of linear logic underpins everything here.
- Lamarche introduced the Cartesian closed category of weighted relations via an algebraic construction

François Lamarche, 'Quantitative Domains and Infinitary Algebras', *Theoretical Computer Science* 94, no. 1 (2 March 1992): 37–62, [https://doi.org/10.1016/0304-3975\(92\)90323-8](https://doi.org/10.1016/0304-3975(92)90323-8).

- Ehrhard's *Finiteness Spaces* refine these models so that infinite sums are not needed.

Thomas Ehrhard, 'Finiteness Spaces', *Mathematical Structures in Computer Science* 15, no. 4 (August 2005): 615–46, <https://doi.org/10.1017/S0960129504004645>.

- Probabilistic coherence spaces were shown to be an adequate model of lambda-calculus by Ehrhard et al.

Thomas Ehrhard, Michele Pagani, and Christine Tasson, 'The Computational Meaning of Probabilistic Coherence Spaces', in *LICS 2011 (IEEE 2011)*, 87–96, <https://doi.org/10.1109/LICS.2011.29>.

We'll describe work that shows a general soundness result for weighted relational models of typed lambda-calculus with recursion.

The simplest model of linear logic

The *relational model* is arguably the simplest categorical model of linear logic:

- A formula A is interpreted as a set $\llbracket A \rrbracket$
- A proof of $A \vdash B$ is interpreted as a relation from $\llbracket A \rrbracket$ to $\llbracket B \rrbracket$
- More generally, a proof of

$$A_1, \dots, A_n \vdash B$$

is interpreted as a subset of

$$\llbracket A_1 \rrbracket \times \dots \times \llbracket A_n \rrbracket \times \llbracket B \rrbracket.$$

The category **Rel** has:

- objects: sets
- morphisms $X \rightarrow Y$: relations from X to Y i.e. subsets of $X \times Y$
- composition: relational composition

- $\llbracket A \otimes B \rrbracket = \llbracket A \multimap B \rrbracket = \llbracket A \rrbracket \times \llbracket B \rrbracket$; cartesian product of sets
- $\llbracket A \& B \rrbracket = \llbracket A \rrbracket + \llbracket B \rrbracket$; disjoint union

Note that

$$\mathbf{Rel}(X \times Y, Z) \cong \mathbf{Rel}(X, Y \times Z).$$

We have a symmetric monoidal closed category, hence a model of the \otimes, \multimap fragment of MLL.

The disjoint union is a Cartesian product in this category, so we can model $\&$.

Exponential

The exponential $!A$ can be interpreted using *finite multisets*:

$$\llbracket !A \rrbracket = \mathbf{M}_f \llbracket A \rrbracket$$

We'll give the comonad structure.

Given a relation $f : X \rightarrow Y$, $!f$ is the relation

$$\{([x_1, \dots, x_n], [y_1, \dots, y_n]) \mid (x_i, y_i) \in f\}.$$

Dereliction (counit) is the relation $!X \rightarrow X$ given by

$$\{([x], x) \mid x \in X\}.$$

Digging (comultiplication) is the relation $!X \rightarrow !!X$ given by

$$\{(m, M) \mid m = \bigsqcup_+ M\}.$$

We also get the Seely isomorphisms: $\mathbf{M}_f(X + Y)$ is isomorphic to $\mathbf{M}_f X \times \mathbf{M}_f Y$.

Could not be easier!

A short abstract interlude

Finite multisets give the *free commutative monoid* on a set.

Because the category **Rel** of sets and relations is the Kleisli category on **Set** for the powerset monad, and because this is commutative, \mathbf{M}_f also gives free commutative monoids in **Rel**.

(See e.g. Hyland et al. 'A Category Theoretic Formulation for Engeler-Style Models of the Untyped λ -Calculus'. Proceedings of MFCSIT 2004, ENTCS 161 (31 August 2006): 43–57. <https://doi.org/10.1016/j.entcs.2006.04.024>.)

Rel is self-dual, so these are also free commutative *comonoids*.

Hence \mathbf{M}_f is the *Lafont exponential* in **Rel**.

A model of typed lambda-calculus

The Kleisli category for \mathbf{M}_f gives us a Cartesian closed category and hence a model of typed lambda calculus.

A term of type $A \rightarrow B$ is interpreted as a relation from $\mathbf{M}_f[[A]]$ to $[[B]]$.

Composition: $f; g$ is given by

$$\{(\bigsqcup_i S_i, n) \mid (S_i, n_i) \in f, ([n_1, \dots, n_k], n) \in g\}$$

$$\begin{array}{ccccccc}
 & & & & n & & \\
 & & & & | & & \\
 & & & & g & & \\
 & & & & | & & \\
 [n_1, & \dots & n_i, & \dots & n_k] & & \\
 f | & & f | & & f | & & \\
 S_1 & \oplus & S_i & \oplus & S_k & &
 \end{array}$$

Assuming an obvious interpretation of the booleans and logical operations, $x \wedge \neg y$ is the following relation:

$x : \mathbf{Bool},$	$y : \mathbf{Bool}$	\vdash	$x \wedge \neg y : \mathbf{Bool}$
[true]	[true]		false
[true]	[false]		true
[false]	[true]		false
[false]	[false]		false

If x and y are contracted to give $x \wedge \neg x$ we have the following relation:

$x : \mathbf{Bool}$	\vdash	$x \wedge \neg x : \mathbf{Bool}$
[true, true]		false
[true, false]		true
[false, true]		false
[false, false]		false

Remarks

It is natural to read these relations right-to-left: “to deliver this output, these inputs are needed”.

The model is already somewhat *quantitative* (or at least *intensional*): it reveals the number of times a term interacts with its variables.

The model caters for nondeterminism: the two invocations of an argument can in principle be supplied *different* values.

A model of nondeterminism

We can exploit this to interpret a nondeterministic extension of the language:

$$\llbracket M \text{ or } N \rrbracket = \llbracket M \rrbracket \cup \llbracket N \rrbracket.$$

From relations to matrices

A relation on $f \subseteq X \times Y$ can be viewed as a matrix indexed by $X \times Y$, populated with booleans:

$$f_{(x,y)} = \text{true} \quad \text{iff} \quad (x, y) \in f.$$

In this view, composition of relations is matrix multiplication: given relations $f : X \rightarrow Y$ and $g : Y \rightarrow Z$

$$(g \circ f)_{(x,z)} = \bigvee_{y \in Y} f_{(x,y)} \wedge g_{(y,z)}.$$

Ingredients:

- finite multiplication (\wedge)
- infinite sum (\bigvee)
- distributivity

Complete semirings

We can replace the booleans in this construction with any *complete semiring*.

A complete (commutative) semiring $(R, \cdot, 1, \sum)$ comprises:

- a commutative monoid $(R, \cdot, 1)$
- summation maps $\sum_X : R^X \rightarrow R$ for every set X satisfying
 - appropriate commutativity and associativity laws
 - distributivity

$$a \cdot \sum_I b_i = \sum_I a \cdot b_i.$$

The empty sum provides a zero element.

Infinite sums

We need infinite sums because some of our indexing sets will be infinite: $\mathbf{M}_f X$ is infinite for any non-empty X .

Ehrhard's *finiteness spaces* give a way to manage without infinite sum: extra structure guarantees that in the composition

$$(g \circ f)_{(x,z)} = \bigvee_{y \in Y} f_{(x,y)} \wedge g_{(y,z)}.$$

there are at most finitely many y for which $f_{(x,y)}$ and $g_{(y,z)}$ are non-zero.

We take the easy way out and assume infinite sums exist.

Complete semirings have no negatives

By the way, a complete semiring cannot be a ring: they are *zerosumfree* (aka *positive*) because if $a + b = 0$ then

$$\begin{aligned} 0 &= 0 + 0 + \dots \\ &= (a + b) + (a + b) + \dots \\ &= a + (b + a) + (b + a) + \dots \\ &= a + 0 + 0 + \dots \\ &= a. \end{aligned}$$

Examples

- $(\mathbf{Bool}, \wedge, \text{true}, \vee)$
 - the booleans with finite conjunction and infinite disjunction
 - this gives us the relational model
- $(\mathbf{N}^\infty, \times, 1, \sum)$
 - the natural numbers with infinite, with finite multiplication and infinite summation
 - \mathbf{N} can be replaced by \mathbf{R}
- $(\mathbf{N}^\infty, +, 0, \min)$
 - the (discrete) *tropical* semiring
- $(\mathbf{R}^\infty, +, 0, \inf)$
 - the (continuous) *tropical* semiring

Categories of *weighted relations*

Given any complete semiring R we can define a category, generalising from **Rel**:

- objects are sets
- morphisms X to Y are matrices which we can describe as functions

$$f : X \times Y \rightarrow R$$

- composition:

$$(f; g)_{(x,z)} = \sum_{y \in Y} f_{(x,y)} \cdot g_{(y,z)}$$

The identity is given by the diagonal matrix,

$$id_{(x,x')} = \delta_{x,x'}$$

using the Kronecker delta.

Weighted relational models of linear logic

It's easy to lift the interpretation of the multiplicatives and $\&$ from **Rel** to a weighted model: nothing changes!

- $\llbracket A \otimes B \rrbracket = \llbracket A \multimap B \rrbracket = \llbracket A \rrbracket \times \llbracket B \rrbracket$
- $\llbracket A \& B \rrbracket = \llbracket A \rrbracket + \llbracket B \rrbracket$

As before, morphisms from $X \times Y$ to Z are the same thing as morphisms from X to $Y \times Z$.

And morphisms from X to $Y + Z$ are matrices

$$f : X \times (Y + Z) \rightarrow R$$

which correspond to pairs of matrices

$$f_1 : X \times Y \rightarrow R \quad f_2 : X \times Z \rightarrow R.$$

So we get a categorical product again.

Algebraic viewpoints

We can view these matrices as describing *linear maps* between the modules of vectors R^X and R^Y ; or in the other direction via duality.

We could also see them as giving, for each $y \in Y$, a vector R^X , which we could write as an infinite formal sum

$$y \mapsto r_1x_1 + r_2x_2 + \dots$$

Exponentials, quantitatively

Exponentials interact more interestingly with the coefficients.

We want to make $\mathbf{M}_f X$ into a (free) commutative comonoid on X ; in the right-to-left view, this will be a monoid.

The underlying set $\mathbf{M}_f X$ is already a monoid, where multiplication is multiset union. We extend this to formal sums (vectors).

Suppose $f : \mathbf{Bool} \rightarrow \mathbf{Bool}$ has

$$f_{(\text{true}, \text{true})} = 1 \quad f_{(\text{false}, \text{true})} = 1$$

In the right-to-left viewpoint outlined above:

$$\text{true} \mapsto 1.\text{true} + 1.\text{false}.$$

Consider $!f : \mathbf{M}_f\mathbf{Bool} \rightarrow \mathbf{M}_f\mathbf{Bool}$. Following algebraic intuition, we define it by multiplication:

$$\begin{aligned} [\text{true}, \text{true}] &\mapsto (1.\text{true} + 1.\text{false}) \times (1.\text{true} + 1.\text{false}) \\ &= 1.[\text{true}, \text{true}] \\ &\quad + (1 + 1). [\text{true}, \text{false}] \\ &\quad + 1.[\text{false}, \text{false}] \end{aligned}$$

Thus $!f : \mathbf{M}_f\mathbf{Bool} \rightarrow \mathbf{M}_f\mathbf{Bool}$ has *two* ways to relate

$$[\text{true}, \text{false}] \quad \text{with} \quad [\text{true}, \text{true}].$$

Kleisli composition is power series substitution

A morphism $!X \rightarrow Y$ associates:

- to every $y \in Y$ a vector of quantities indexed by $\mathbf{M}_f X$.

If we regard the elements of x as formal indeterminates, this is:

- to every $y \in Y$ a formal power series

$$\sum_{m \in \mathbf{M}_f X} f_y(m) \cdot m$$

Each multiset is a basic monomial, something like

$$x_1^3 x_2 x_3^2 = [x_1, x_1, x_1, x_2, x_3, x_3]$$

and the quantities are the coefficients.

Given $f : \mathbf{M}_f X \rightarrow Y$, we want the promotion $f^! : \mathbf{M}_f X \rightarrow \mathbf{M}_f Y$ to be a *comonoid homomorphism*. With the above view, this means

- to every multiset (monomial) in $\mathbf{M}_f Y$, associate a formal power series
- in such a way that *power series multiplication* is preserved.

The functorial action is similar:

- $f : X \rightarrow Y$ associates to each y a coefficient for each x , i.e. a formal sum of $\$x\s
- we lift this to $!f : \mathbf{M}_f X \rightarrow \mathbf{M}_f Y$ by multiplication of power series

Given $f : \mathbf{M}_f X \rightarrow Y$ and $g : \mathbf{M}_f Y \rightarrow Z$, the Kleisli composition $f; g$ works as follows:

- g takes each z to a power series over Y ;
- substitute each $y \in Y$ with the power series $f(y)$ and multiply out.

Example

Let $f, g : \mathbf{M}_f \mathbf{Bool} \rightarrow \mathbf{Bool}$ be given by:

$$g(\text{true}) = 1.\text{true}^2 + 2.\text{false}$$

$$g(\text{false}) = 0$$

$$f(\text{true}) = 1.\text{true} + 1.\text{false}$$

$$f(\text{false}) = 1.\text{true}$$

Then $f; g$ maps true to

$$(1.\text{true} + 1.\text{false})^2 + 2.\text{true} = 1.\text{true}^2 + 2.\text{truefalse} + 1.\text{false}^2 + 2.\text{true}$$

Abstractly

What we've constructed is the *symmetric algebra* (free commutative monoid) over the module of X -indexed vectors.

But we're reading our arrows backwards, so this is the free commutative comonoid: the Lafont exponential.

Lamarche made the observation that

Linear logic is in fact co-linear

It's quite common to discover models of linear logic (and related things) as the *opposite* of familiar algebraic categories.

Weighted relational models of typed lambda-calculi

These constructions give us a model of linear logic, and the Kleisli category is Cartesian closed.

We can therefore model typed lambda calculi. Let's model something like a programming language:

- add some base types **Bool**, **Int**
- add some constants with the obvious semantics e.g. $\llbracket \text{true} \rrbracket : \mathbf{M}_f \emptyset \rightarrow \mathbf{Bool}$ has

$$\llbracket \text{true} \rrbracket_{([\], \text{true})} = 1 \quad \llbracket \text{true} \rrbracket_{([\], \text{false})} = 0$$

We have additional operations available in the model:

- we can add weighted relations together
- we can multiply relations by quantities.

What does this let us do?

Nondeterminism

Addition of morphisms gives us an interpretation of nondeterministic choice:

$$\llbracket M \text{ or } N \rrbracket = \llbracket M \rrbracket + \llbracket N \rrbracket.$$

If we work with the semiring $(\mathbf{N}^\infty, \times, 1, \sum)$, we see that for example

$$\llbracket \text{true or true} \rrbracket_{([], \text{true})} = 2.$$

More generally we might expect that for a term $\vdash M : \mathbf{Bool}$,

$$\llbracket M \rrbracket_{([], \text{true})}$$

is the number of choice-paths that M can take to reach `true`.

Probabilistic choice

Since we can *weight* our terms by multiplication with a quantity, we can interpret *probabilistic* choice: if we work with the reals rather than naturals we can define

$$\llbracket M +_p N \rrbracket = p\llbracket M \rrbracket + (1 - p)\llbracket N \rrbracket.$$

Then for example

$$\llbracket \text{true} +_{\frac{1}{2}} (\text{true} +_{\frac{1}{2}} \text{false}) \rrbracket_{(\[], \text{true})} = \frac{3}{4}.$$

Can we use $[0, 1]$?

For modelling probability, it's tempting to try using coefficients from $[0, 1]$ rather than \mathbf{R}^∞ .

Sadly this won't work. For example, the term

$$x : \mathbf{Bool}, y : \mathbf{Bool} \vdash x \wedge y : \mathbf{Bool}$$

contains

$([\mathbf{true}], [\mathbf{false}], \mathbf{false})$ with coefficient 1

$([\mathbf{false}], [\mathbf{true}], \mathbf{false})$ with coefficient 1

If we contract x and y , the input patterns

$([\text{true}], [\text{false}])$ $([\text{false}], [\text{true}])$

become the same multiset and the coefficients are *added*.

Then

$x : \mathbf{Bool} \vdash x \wedge x : \mathbf{Bool}$

contains

$([\text{true}, \text{false}], \text{false})$ with coefficient 2

The good news is that the model still seems to be counting choice-paths correctly!

A generic soundness theorem

We shall give a weighted operational semantics to typed lambda-calculus in such a way that these relational models measure the operational reduction paths.

Syntax (at least):

$$M ::= x \mid \lambda x^A. M \mid MM \mid M \text{ or } M \mid rM$$

for $r \in R$, the semiring of interest; plus base types and constants for booleans, numerals and basic operations as desired.

Operational semantics is via *weighted reductions* $M \xrightarrow{r} N$: “ M reduces to N in one step with weight r ”.

Basic reductions

$$(\lambda x^A. M)N \xrightarrow{1} M[N/x]$$

$$M \text{ or } N \xrightarrow{1} M \quad M \text{ or } N \xrightarrow{1} N$$

$$rM \xrightarrow{r} M$$

Evaluate in a call-by-name strategy.

Define the *weight* of a reduction path

$$M \xrightarrow{r_1} M_1 \cdots \xrightarrow{r_n} M_n$$

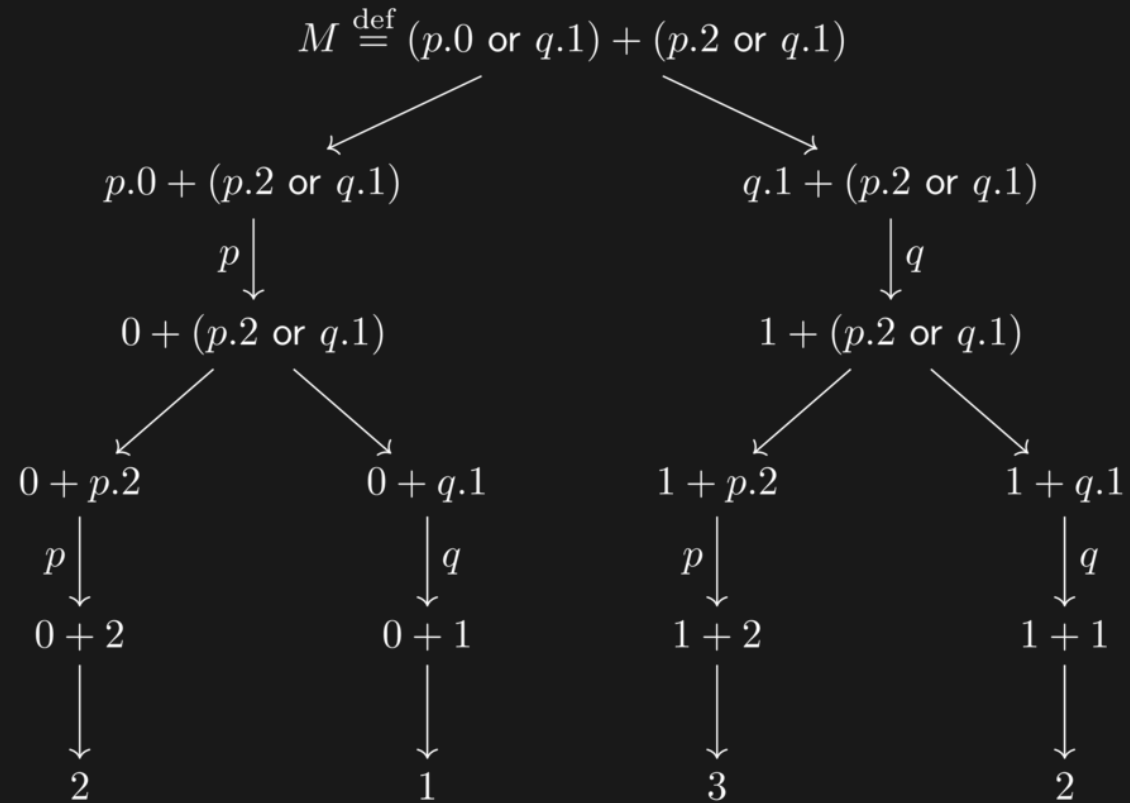
to be the product

$$r_1 \cdot \cdots \cdot r_n$$

For a term $\vdash M : \mathbf{Int}$, define $\text{Red}(M, n)$ to be the sum of the weights of paths

$$M \rightarrow^* n.$$

Example



$$\text{Red}(M, 2) = p^2 + q^2$$

Soundness

By standard reasoning we can prove

Theorem. For any term $\vdash M : \mathbf{Int}$

$$\mathbf{Red}(M, 0) = \llbracket M \rrbracket_{([], 0)}.$$

- Using $(\mathbf{N}^\infty, \times, 1, \sum)$, the semantics counts reduction paths
- Using $(\mathbf{R}^\infty, \times, 1, \sum)$, translating probabilistic choice as

$$\frac{1}{2}M \text{ or } \frac{1}{2}N$$

the semantics computes probability of reduction to a value

- Using $(\mathbf{N}^\infty, +, 0, \min)$, the model computes the lowest weight of a reduction path;
 - if we instrument terms, e.g. replacing each $\lambda x. M$ by $\lambda x. 1 M$, the model counts β -reductions.

Fixed points and recursion

To interpret recursion, we'd like to add a fixed point combinator to the language and give it a semantics.

The easy way: insist that our semiring is *continuous*.

That is:

- the semiring is a complete partial order
- addition and multiplication are continuous
- infinite sums are least upper bounds of chains of finite sums.

Then the morphisms form complete partial orders and we can interpret fixed points in the normal way. Soundness is proved by standard techniques.

Fixed points without continuity

In fact, as shown by Laird, continuity is not required.

A fixed point combinator may be defined by taking a sum over all possible recursive call-patterns, expressed using *nested multisets*:

$$\mathbb{M}_0 = \emptyset \quad \mathbb{M}_{n+1} = \mathbf{M}_f \mathbb{M}_n \quad \mathbb{M} = \bigcup_n \mathbb{M}_n.$$

Given $f : \mathbf{M}_f X \rightarrow X$, we can define $f^m : I \rightarrow \mathbf{M}_f X$ for each $m \in \mathbb{M}$ expressing the result of nested recursive calls to f following the pattern of m .

Then

$$\text{fix}(f) = \sum_{m \in \mathbb{M}} f^m; f.$$

J. Laird, 'Fixed Points In Quantitative Semantics', in Proceedings, LICS '16, ACM 2016), 347–56, <https://doi.org/10.1145/2933575.2934569>.

Untyped lambda-calculus

A related construction gives a model of untyped λ -calculus in these categories:

$$D_0 = \emptyset \quad D_{n+1} = (\mathbf{M}_f D_n)^\omega \quad D = \bigcup D_n.$$

Then $D \cong \mathbf{M}_f D \times D$ so we have a reflexive object

$$D \cong D \Rightarrow D.$$

Antonio Bucciarelli, Thomas Ehrhard, and Giulio Manzonetto, 'Not Enough Points Is Enough', in CSL 2007, Springer 2007), 298–312, https://doi.org/10.1007/978-3-540-74915-8_24.

Further

Our simple-minded quantitative model is *sound* but not *fully abstract* and has barely been explored. Others have done more:

- Probabilistic coherence spaces are a refinement of one version of this model that delivers a quite remarkable full abstraction result.

Thomas Ehrhard, Christine Tasson, and Michele Pagani, 'Probabilistic Coherence Spaces Are Fully Abstract for Probabilistic PCF', in Proceedings, POPL '14 (ACM 2014), 309–20, <https://doi.org/10.1145/2535838.2535865>.

- Barbarossa and Pistone have recently studied the tropical version of the model. They have established connections with generalized metric spaces, and analysed the differential structure of the model in detail.

Davide Barbarossa and Paolo Pistone, 'Tropical Mathematics and the Lambda-Calculus I: Metric and Differential Analysis of Effectful Programs', in CSL 2024, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2024, <https://doi.org/10.4230/LIPIcs.CSL.2024.14>.