Normal funct~~ors~~[ions], [the irrelevence of] power series, and [a new model of] λ-calculus.

MORGAN ROGERS, Université Sorbonne Paris Nord, France
THOMAS SEILLER*, CNRS, France
WILLIAM TOIANI, Melbourne University, Australia

1 INTRODUCTION

The language of mathematics is constrained by the finite means we have to express it. Simultaneously, we are interested in inherently infinite structures. Thus, the language we use to describe mathematics must thread a needle between the finite and the infinite. Modern mathematical foundations achieves this with great success; we indulge ourselves in the hypothesis that a new variable may always be introduced and distinguished from the finite set which are currently in use. This is instilled in the assumption that an infinite set (usually taken to be countable) of variables exists, yet any individual mathematical proof contains only a finite subset of these variables. In this way, the variables used inside mathematical proof are merely *potentially infinite*.

The untyped λ-calculus reflects this "potential infinitude": in the context of a redex $(\lambda x.M)N$ there are an *arbitrary* yet *finite* number of free occurrences of x inside M, and thus an arbitrary yet finite number of substitutions performed in the single-step β-reduction $(\lambda x.M)N \longrightarrow_\beta M[x := N]$. This also holds for the simply typed λ-calculus, however the *untyped* λ-calculus admits another dimension of potential infinitude in that there are terms whose reductions grow *arbitrarily large*, although each term itself in the reduction sequence is finite. For instance we have, where ω denotes $\lambda x.xxx$:

$$\omega\omega \longrightarrow_\beta \omega\omega\omega \longrightarrow_\beta \omega\omega\omega\omega \longrightarrow_\beta \dots$$

The current paper concerns itself with models of the untyped λ-calculus. To illustrate how a model can capture potential infinitude, consider a set A and let $\mathcal{P}(A)$ denote the powerset of A. Observe that any subset $X \in \mathcal{P}(A)$ can be written as the filtered colimit of finite sets $\operatorname{colim}_{X' \subseteq_{\text{fin}} X} X'$. Suppose B is another set and we have a function $f : A \longrightarrow B$. Then f induces a map $f_* : \mathcal{P}(A) \to \mathcal{P}(B)$ which preserves (amongst other things) filtered colimits and wide pullbacks, so that we may evaluate f_* at a given subset X as the union of f_* applied to the finite subsets of X. We can think of the collection $\{f_*(X') \mid X' \subseteq_{\text{fin}} X\}$ as a collection of *finite approximations* to f_*. Moreover, if $X', X'' \subseteq A$ are both finite such that $x \in X', X''$ then $x \in X' \cap X''$ and so

$$f_*(X' \cap X'') = f_*(X') \cap f_*(X'') \tag{1}$$

as f_* preserves pullbacks. This implies that there exists a *minimal, finite* subset $\hat{X}$ determining the behaviour of f on some $x \in A$.

The phenomenon just observed lifts to a more general setting. We replace finite sets with finite functors (presheaves F such that $\bigsqcup_{a \in A} F(a)$ is finite) along with a canonical choice of representative for each finite set $F(a)$. For the set with n elements we take the choice $\{0, \dots, n-1\}$ and the set with 0 elements is the is the unique such, ie, the empty set $\varnothing$.

*T. Seiller is partially supported by the DIM RFSI Exploratory project CoHOp, and the ANR ANR-22-CE48-0003-01 project DySCo.

Authors' addresses: Morgan Rogers, LIPN – UMR 7030 CNRS, Université Sorbonne Paris Nord, 99, avenue Jean-Baptiste Clément, Villetaneuse, 93430, France, rogers@lipn.univ-paris13.fr; Thomas Seiller, LIPN – UMR 7030 CNRS & Sorbonne Paris Nord University, CNRS, 99, avenue Jean-Baptiste Clément, Villetaneuse, 93430, France, seiller@lipn.fr; William Toiani, Melbourne University, Melbourne University, Melbourne, Australia, .

Let $\mathrm{Int}(A)$ be the subcategory of Set^A on the *integral functors* (finite functors for which $F(a)$ is one of the above canonical choices, for each $a \in A$).

Lemma 1. *Let A be a set viewed as a category. Any functor $F \in \mathrm{Set}^A$ is the colimit of functors in* $\mathrm{Int}(A)$.

Moreover, *normal functors* (those preserving filtered colimits and wide pullbacks) $\mathscr{F} : \mathrm{Set}^A \longrightarrow \mathrm{Set}$ can be represented using finite data.

Motivated by a search for a satisfying model of the untyped λ-calculus, Girard succeeded in defining a model where each λ-term t is interpreted with respect to a finite context $\{x_1, \ldots, x_n\}$ as a normal functor

$$\mathrm{Set}^A \times \ldots \times \mathrm{Set}^A \longrightarrow \mathrm{Set}^A$$

where there are n copies of Set^A in the domain.

The key technical tool used in Girard's proof is the Normal Form Theorem, which equates normal functors to *analytic functors*, functors admitting a presentation reminiscent of formal power series. However, faced with the complexity of the details of this model, one must wonder: *is the categorical framework of presheaves absolutely necessary?* Or in other words, is there a simpler structure underlying this categorical framework which is still sufficient to model the λ-calculus?

2 A SIMPLE COMBINATORIAL MODEL

Our starting point was to replace integral functors $F \in \mathrm{Int}(A)$ with finite multisets of elements of A; we denote the set of such things by $\mathcal{I}(A)$. This embeds inside the set of functions $A \longrightarrow \overline{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$; we denote the set of such functions by $\mathcal{Q}(A)$, which replaces Set^A. These sets come with a natural pointwise ordering, a structure much simpler than the natural transformations in Set^A.

In this context, we found a new model where a term t in context $\underline{x} = \{x_1, \ldots, x_n\}$ is interpreted as a *normal function*: an order preserving function

$$[\![\underline{x} \mid t]\!] : \mathcal{Q}(A) \times \ldots \times \mathcal{Q}(A) \longrightarrow \mathcal{Q}(A) \tag{2}$$

which preserves directed suprema.

Definition 2. An order preserving function $f : \mathcal{Q}(A)^n \longrightarrow \mathcal{Q}(B)$ is **analytic** if for any pair $(x, b) \in \mathcal{Q}(A)^n \times B$ we have

$$f(x)(b) = \sup\nolimits_{u \in \mathcal{I}(A)^n} f(u)(b)\delta_{u \leq x} \tag{3}$$

where

$$\delta_{u \leq x} = \begin{cases} 1, & u \leq x \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Theorem 3. *Let $f : \mathcal{Q}(A)^n \longrightarrow \mathcal{Q}(B)$ be order preserving. Then f is normal if and only if it is analytic.*

Currying is realised in this model as a bijection between normal functions $\mathcal{Q}(A)^n \times \mathcal{Q}(B) \longrightarrow \mathcal{Q}(C)$ and normal functions $\mathcal{Q}(A)^n \longrightarrow \mathcal{Q}(\mathcal{I}(B) \times C)$.

In order to reach a full model of the *untyped* lambda calculus, we fix a set A which is countably infinite. This assumption allows us to construct a bijection

$$q : \mathcal{I}(A) \times A \longrightarrow A. \tag{5}$$

There is an induced bijection

$$\begin{aligned} \overline{q} : \mathcal{Q}(A) &\longrightarrow \mathcal{Q}(\mathcal{I}(A) \times A) \\ \underline{a} &\longmapsto \underline{a} \circ q. \end{aligned}$$

Speaking loosely, the bijection $\overline{q}$ allows a passage beginning at a normal function $f : \mathcal{Q}(A)^n \times \mathcal{Q}(A) \longrightarrow \mathcal{Q}(A)$ which can be curried to $f^+ : \mathcal{Q}(A)^n \longrightarrow \mathcal{Q}(\mathcal{I}(A) \times A)$ and then composed with $\overline{q}^{-1}$ to obtain a function $\overline{q}^{-1} f^+ : \mathcal{Q}(A)^n \longrightarrow \mathcal{Q}(A)$. This is how abstraction is modelled.

Lemma 4 (Substitution Lemma)**.** *Let t, s be λ-terms and $\underline{x} = \{x_1, \ldots, x_n\}$ be a collection of variables and y another variable so that $\underline{x} \cup \{y\}$ is a valid context for t and $\underline{x}$ is a valid context for s. Then for any $\alpha \in \mathcal{Q}(A)^n$ we have*

$$[\![\underline{x} \mid t[y := s]]\!](\alpha) = [\![\underline{x}, y \mid t]\!](\alpha, [\![\underline{x} \mid s]\!](\alpha)) \tag{6}$$

Theorem 5. *This is a denotational model of the λ-calculus. That is, if t is a λ-term and $\underline{x}$ a valid context for t and for s, then we have the following equality.*

$$[\![\underline{x} \mid (\lambda y.t)s]\!] = [\![\underline{x} \mid t[y := s]]\!] \tag{7}$$

Having found this "decategorification" of Girard's original work, we can "recategorify": extract the essential features to further clarify 'why' this model works and generate further models sharing these features. Indeed, having reduced to functions into $\overline{\mathbb{N}}$, we could further consider analogous functions into other sets. However, before embarking on this, we note that Girard *also* decategorified his model, but in a different way. Where we have switched from functors to order preserving functions, Girard switched the codomain of the functors from Set to the two point category $\{0, 1\}$ (admitting only identity morphisms). This lead to coherence spaces which Girard now usually claims to be the origin of Linear Logic [2].

3 LINEAR LOGIC

What does one obtain from this model of λ-calculus? One observation is that analytic functions $\mathscr{F} : \mathcal{Q}(A) \times \cdots \times \mathcal{Q}(A) \longrightarrow \mathcal{Q}(A)$, by (3), are determined by their restriction to the domain $\mathcal{I}(A) \times \ldots \times \mathcal{I}(A)$. If we think of the untyped λ-calculus as a *typed* system, where every term is of the same type A, it appears as though $\mathcal{I}$ is modeling some unnoticed connective. Can our model be decomposed to a semantics of a finer syntactic system? The "finer" granility will need to include functions $\mathcal{Q}(A) \times \cdots \times \mathcal{Q}(A) \longrightarrow \mathcal{Q}(A)$ which are determined by their restrictions to the domain $A \times \cdots \times A$, and these are the additive functions!

Definition 6. Given sets $A_1, \ldots, A_n, B$, a function $f : \mathcal{Q}(A_1) \times \ldots \times \mathcal{Q}(A_n) \longrightarrow \mathcal{Q}(B)$ is *additive* if it is linear in each argument. We denote the set of all additive functions

$$\mathrm{App}(\mathcal{Q}(A_1) \times \ldots \times \mathcal{Q}(A_n), \mathcal{Q}(B)) \tag{8}$$

Lemma 7. *Let $A_1, \ldots, A_{n+1}, B_1, \ldots, B_m$ be sets. There is a bijection*

$$\begin{aligned} &\mathrm{App}\left(\mathcal{Q}(A_1) \times \ldots \times \mathcal{Q}(A_n) \times \mathcal{Q}(A_{n+1}), \mathcal{Q}(B)\right) \\ &\qquad \longrightarrow \mathrm{App}\left(\mathcal{Q}(A_1) \times \ldots \times \mathcal{Q}(A_n), \mathcal{Q}(A_{n+1} \times B)\right) \end{aligned}$$

For example, if $f : \mathcal{Q}(A_1) \times \ldots \times \mathcal{Q}(A_n) \times \mathcal{Q}(A_{n+1}) \longrightarrow \mathcal{Q}(B)$ is additive, $\underline{a}_i \in \mathcal{Q}(A_i)$ for $i = 1, \ldots, n+1$, and $b \in B$ is arbitrary, then

$$\begin{aligned} &f(\underline{a}_1, \ldots, \underline{a}_n, \underline{a}_{n+1})(b) \\ &= \sum_{c_1, \ldots, c_{n+1} \in A_1, \ldots, A_{n+1}} \underline{a}_1(c_1) \ldots \underline{a}_{n+1}(c_{n+1}) f(\iota_1(c_1), \ldots, \iota_n(c_n), \iota_{n+1}(c_{n+1}))(b) \end{aligned}$$

We define the map $f^\times : \mathcal{Q}(A_1) \times \ldots \mathcal{Q}(A_n) \longrightarrow \mathcal{Q}(A_{n+1} \times B)$ as follows.

$$\begin{aligned} &f^\times(\underline{a}_1, \ldots, \underline{a}_n)(c_{n+1}, b) \\ &= \sum_{c_1, \ldots, c_n \in A_1, \ldots, A_n} \underline{a}_1(c_1) \ldots \underline{a}_{n+1}(c_{n+1}) f(\iota_1(c_1), \ldots, \iota_{n+1}(c_{n+1}))(b) \end{aligned}$$

This is how we anticipate (R $\multimap$) in sequential linear logic to be modeled:

$$\dfrac{\begin{array}{c}\pi' \\ \vdots \\ A_1, \ldots, A_n, A_{n+1} \vdash B_1, \ldots, B_m\end{array}}{A_1, \ldots, A_n \vdash \neg A_{n+1}, B_1, \ldots, B_m}\ (\text{R} \multimap)$$

we take

$$[\![\pi]\!] = [\![\pi']\!]^{\times} : Q(A_1) \times \ldots \times Q(A_n) \longrightarrow Q(A_{n+1} \times B_1 \times \ldots \times B_m)$$

4 RECATEGORIFYING

The remainder of our work consists of a more technical contribution: the categorification of our model to obtain a class of models for linear logic.

First of all, we observe that if we view all of the functions in the above discussion as functors, then Int and $\mathcal{I}$ are functors Set $\to$ CAT of a rather specific form; namely, they are the class of functors 'of finite support' into a category. In order for the notion of finite support to make sense, we need a category C having a distinguished zero object 0, so that a finite support functor $X \to C$ is one which is equal to 0 at all but finitely many elements. We write

$$\mathcal{I}(X) := \mathrm{Fun}_{fs}(X, C)$$

for the collection of such functors. In order for $\mathcal{I}$ to be functorial we need to include a 'monoidal sum' $\oplus$ for which 0 is the unit; then a function $X \to Y$ induces a function $\mathcal{I}(X) \to \mathcal{I}(Y)$ via

$$\mathcal{I}(f)(F)(y) := \bigoplus_{x \in f^{-1}(y)} F(x),$$

which is well-defined since there are only finitely many x for which $F(x)$ is not 0.

Remark 8. Since $\mathcal{I}$ is a small category, the same is true of $\mathcal{I}(X)$ and we can compose with the functor returning the set of (isomorphism classes of) objects in order to view it as an endofunctor on Set. However, when we introduce Q below, we encounter size issues (at least in the original case of Set^X considered by Girard). For simplicity, we will treat $\mathcal{I}$ and Q as endofunctors and monads on Set in the present text. For a formal treatment which avoids the large cardinality assumptions as Girard's model was supposed to, we instead require *relative pseudomonads*.

In order for $\mathcal{I}$ to be a monad, we need a monoidal product $\otimes$ on C with a unit object 1 such that $\otimes$ distributes over $\oplus$. This enables us to reproduce the 'singleton' functor δ_x which is 1 at x and 0 at $x' \neq x$ as the unit of the monad, and with multiplication

$$\begin{aligned} \mu : \mathcal{I}(\mathcal{I}(X)) &\to \mathcal{I}(X) \\ (G : \mathcal{I}(X) \to C) &\mapsto \left(x \mapsto \bigoplus_{F \in \mathcal{I}(X)} G(F) \otimes F(x) \right), \end{aligned}$$

where once again the monoidal sum is well defined since G has finite support. The associativity axiom for this multiplication map corresponds to the requirement that $\otimes$ distributes over $\oplus$. In summary, the structure needed is that of a **rig category**, a category equipped with two unital monoidal operations such that one distributes over the other.

Next, the general situation is which a class of structures is generated by its subclass of finite (or finitely presented) structures is captured by the *ind-completion* ('ind' was originally an abbreviation of 'inductive'). The ind-completion Ind(C) of a category C is its free cocompletion under directed colimits. This construction arises frequently in logic, and especially in universal algebra, because the class of models of an algebraic theory is always the ind-completion of the class of finitely presentable models of that theory.

For a rig category C, $\mathrm{Ind}(C)$ inherits the structure of a rig category. It turns out that if the monoidal sum $\oplus$ is **semi-cartesian**, meaning that 0 is an initial object, then $\mathrm{Ind}(C)$ actually acquires infinite monoidal sums. With this added assumption, we can define

$$Q(X) := \mathrm{Fun}(X, \mathrm{Ind}(C)),$$

which is a relative pseudomonad with the extensions of the same definitions as above. The most interesting part of this 'recategorification' is how we use relative pseudomonads, whereas the standard story is to use a single comonad.

5 BICATEGORICAL SEMANTICS

Regular categorical models interpret cut-equivalent proofs as the same object to equality. Models where two such proofs are identified only up to isomorphism in the model have been of recent interest to many modern developments of the material presented here. However, our goal is to simplify the categorical model given in Girard's original paper and then to find an appropriate categorification of this simplification. Indeed, we anticipate all equalities to be strict, and so the bicategorical semantics research regime is only peripherally related to our work, and is not directly relevant.

REFERENCES

[1] *Sense and Reference* G. Frege. Philosophical Review 57 (3):209-230 (1948)
[2] J.Y.Girard, *The Blind Spot.* European Mathematical Society, 2011.
[3] *Linear Logic*, J.Y. Girard. Theoretical Computer Science, Volume 50, Issue 1, Jan. 30, 1987.
[4] J. Y. Girard, *Normal functors, power series and lambda-calculus.* Annals of Pure and Applied Logic Volume 37, Issue 2, February 1988, Pages 129-177
[5] Hasegawa, *Two applications of analytic functors.* Theoretical Computer Science 272 (2002) 113–175.
[6] *Towards a Typed Geometry of Interaction* CSL 2005: Computer Science Logic pp 216–231.
[7] *Linear Logic and the Hilbert Space* Advances in Linear Logic , pp. 307 - 328, Cambridge University Press, 1995.
[8] *Lectures on the Curry-Howard Isomorphism* Published: July 4, 2006 Imprint: Elsevier Science