Tropical Mathematics and the Lambda-Calculus

Davide Barbarossa  
DISI, Università di Bologna  
davide.barbarossa@unibo.it

Paolo Pistone  
DISI, Università di Bologna  
paolo.pistone2@unibo.it

In recent years, more and more interest in the programming language community has been directed towards the study of quantitative properties of programs like computing the number of computation steps or convergence probabilities, as opposed to purely qualitative properties like termination or program equivalence. In particular, two different quantitative approaches have received considerable attention from the programming language community. On the one hand, the approach of program metrics [2,3,30] and quantitative equational theories [25] is based on the observation that probabilistic or numerical algorithms are not thought to compute a target function \( f \) exactly, but only in an approximate way. This led to study denotational frameworks in which types are endowed with metrics measuring similarities in program behavior [30, 4, 9, 15, 29]. On the other hand, there is the approach based on differential [13, 13, 1, 7, 19] or resource-aware [6] extensions of the \( \lambda \)-calculus, which is well-connected to the relational semantics [12, 19, 23] and non-idempotent intersection types [10, 26]. This led to study syntactic or denotational frameworks in which one can define a Taylor expansion of programs.

In both approaches a crucial role is played by the notion of linearity, in the sense of linear logic, i.e. of using inputs exactly once. In metric semantics, linear programs correspond to non-expansive functions, i.e. maps that do not increase distances; moreover, the possibility of duplicating inputs leads to interpret programs with a fixed duplication bound as Lipschitz-continuous maps [2]. By contrast, in the standard semantics of the differential \( \lambda \)-calculus, linear programs correspond to linear maps, in the usual algebraic sense, while the possibility of duplicating inputs gives rise to power series.

The starting observation of this work is that, at a first glance, there seems to be a “logarithmic” gap between the two approaches: in metric models \( n \) times duplication results in a \( n \)-Lipschitz linear function \( n \cdot x \), while in differential models this results in a non-Lipschitz polynomial function \( x^n \). At the same time, this gap may be overcome once we interpret these functions in the framework of tropical mathematics where, for instance, \( x^n \) precisely reads as \( n \cdot x \).

Tropical mathematics [31] is a well established algebraic and geometrical framework, with tight connections with optimisation theory [22], where the usual ring structure of numbers based on addition and multiplication is replaced by the semiring structure given, respectively, by “\( \min \)” and “\( + \)”’. For instance, the polynomial \( p(x,y) = x^2 + xy^2 + y^3 \), when interpreted over the tropical semiring, translates as the piecewise linear function \( tf(x,y) = \min\{2x, x + 2y, 3y\} \).

A tropical variant of relational semantics has already been considered [19], and shown capable of capturing best-case quantitative properties. Connections between tropical linear algebra and metric spaces have also been observed [14] within the abstract setting of quantale-enriched categories [17,33].

However, a thorough investigation of the full power of the interpretation of the \( \lambda \)-calculus within tropical mathematics has not yet been undertaken. We sketch here some first steps. The aim is to bridge the two approaches mentioned above by making them coexist, and suggesting the application of tropical methods to the study of the \( \lambda \)-calculus and its quantitative extensions. This also scales to a more abstract level, leading to introduce a differential operator for continuous functors between generalized metric spaces (in the sense of [20]).
1 The Tropical Semantics of Linear Logic

Tropical mathematics in a nutshell. We let the tropical semiring \( \mathbb{L} \), the structure at the heart of tropical mathematics, be \([0, \infty)\) with addition \( \min \) and multiplication \( + \). This coincides with the Lawvere quantale \( \mathbb{L} \cong [1, \infty] \), i.e., \([0, \infty)\) with order \( \geq \) and usual \( + \) as the monoid action. \( \mathbb{L} \) is at the heart of the categorical study of metric spaces initiated by Lawvere [20], a viewpoint we will take in the last section. A tropical polynomial is a piece-wise linear function \( \varphi : \mathbb{L} \to \mathbb{L} \) of the form \( \varphi(x) = \min_{i=1}^{\infty} \{i x + \tilde{\varphi}_i\} \), with \( \tilde{\varphi}_i \in \mathbb{N} \). Those are always Lipschitz functions. For example, \( \varphi_n(x) = \min_{i \leq n} \{ix + 2^{-i}\} \), plotted in Fig. 1. A tropical root of \( \varphi \) is a point \( x \in \mathbb{L} \) where the minimum defining \( \varphi \) is attained at least twice. For example, the tropical roots of \( \varphi_{n+1} \) are of the form \( 2^{-(i+1)} \), \( i \leq n \). A tropical Laurent series (of one variable \( x \in \mathbb{L} \)), shortly a tLS, is a function \( \varphi : \mathbb{L} \to \mathbb{L} \) of the form \( \varphi(x) = \inf_{n \in \mathbb{N}} \{nx + \tilde{\varphi}_n\} \), with \( \tilde{\varphi}_n \in \mathbb{L} \). That is, a tLS is a “limit” of tropical polynomials of higher and higher degree. For example \( \varphi(x) := \inf_{n \in \mathbb{N}} \{ix + 2^{-i}\} \) is the “limit” of the \( \varphi_n \), see Fig. 1. Finally, for a polynomial/power series \( f(x) = \sum a_n x^n \), one defines its tropicalization \( tf(\alpha) := \inf_{n \in \mathbb{N}} \{- \log a_n + n \alpha\} \). tLS are in general not Lipschitz, and their study is less developed than that of tropical polynomials.

Tropical weighted relational semantics in a nutshell. The study of matrices with values over the tropical semiring is a special case of the weighted relational semantics [19], a well-studied semantics of the \( \lambda \)-calculus and linear logic: for a fixed continuous semi-ring \( \mathbb{Q} \), take the category \( \mathbb{Q}\text{Rel} \) whose objects are sets and \( \mathbb{Q}\text{Rel}(X, Y) = \mathbb{Q}^{X \times Y} \) (set-indexed matrices with coefficients in \( \mathbb{Q} \)). As expected, \( \mathbb{Q}^X \) is a \( \mathbb{Q} \)-module and we can identify \( \mathbb{Q}\text{Rel}(X, Y) \) with the set of linear maps from \( \mathbb{Q}^X \) to \( \mathbb{Q}^Y \). Taking \( \mathbb{Q} := \mathbb{L} \) we obtain the tropical weighted relational model \( \mathbb{L}\text{Rel} \). Remark that the composition in \( \mathbb{L}\text{Rel} \) reads as \((s \circ t)_{a, c} := \inf_{b \in Y} \{s_{b, c} + t_{a, b}\} \); similarly, linear maps \( f : \mathbb{L}^X \to \mathbb{L}^Y \) are of shape \( f(x)_b = \inf_{a \in X} \{x_a + \tilde{f}_{a, b}\} \), for some matrix \( \tilde{f} \in \mathbb{Q}^{X \times Y} \) and precisely those induced by \( \tilde{f} \in \mathbb{L}\text{Rel}(X, Y) \). We call them tropical linear. By applying known results (taken from [19], [18], [21]), one obtains that \( \mathbb{L}\text{Rel} \) gives rise to denotational models of several variants of the simply typed \( \lambda \)-calculus (STLC): first, \( \mathbb{L}\text{Rel} \) is a SMCC, i.e. a model of the linear STLC. Also, the coKleisli \( \mathbb{L}\text{Rel}_l \) is CCC, i.e. a model of STLC, where \( ! \) is the usual multiset comonad (so \( !X \) is the set of finite multisets on \( X \)). Here, the coKleisli composition of \( s \in \mathbb{L}^{Y \times Z} \) and \( t \in \mathbb{L}^{X \times Y} \) is the matrix \( s \circ t \in \mathbb{L}^{X \times Z} \) given by \((s \circ t)_{\mu, e} := \inf_{n \in \mathbb{N}, b_1, \ldots, b_n \in Y, \mu = \mu_1 + \cdots + \mu_n} \{s_{[b_1, \ldots, b_n], e} + \sum_{i=1}^{n} t_{\mu_i, b_i}\} \). As usual, a matrix \( t \in \mathbb{L}\text{Rel}_l(X, Y) \) yields a linear map \( \mathbb{L}^X \to \mathbb{L}^Y \). However, we can also “express it in the base \( X \)” and see it as a non-linear map \( t^l : \mathbb{L}^X \to \mathbb{L}^Y \) by setting \( t^l(x) := t \circ x \). Concretely, we have \( t^l(x)_b = \inf_{\mu \in \mathbb{L}^X} \{\mu x + t_{\mu, b}\} \). These functions correspond then to generalised tLS, i.e. with possibly infinitely many variables (as many as the elements of \( X \)) and for \( X = Y = \{\ast\} \), we get usual tLS of one variable. Instead, if the support \( \{\mu \in \mathbb{L}^X \mid t_{\mu, b} \neq \infty\} \) of \( f \) is finite, we get tropical polynomials in possibly infinitely many variables. Furthermore, \( ! \) can be decomposed into a family of graded exponentials \( \lambda x \in \mathbb{L} \) turning \( \mathbb{L}\text{Rel}_l \) into a model for a simply typed \( \lambda \)-calculus with bounded duplications, on the style of [8], call it bSTLC.

Finally, \( \mathbb{L}\text{Rel}_l \) can be equipped with a differential operator \( D : \mathbb{L}\text{Rel}_l(X, Y) \to \mathbb{L}\text{Rel}_l(\!(X + X), Y) \) defined by: \( (D \mu)_{x, p, b} := t_{p, \mu, b} \) if \#\( \mu = 1 \) and := \( \infty \) otherwise. This turns \( \mathbb{L}\text{Rel}_l, D \) into a CCC\( \partial C \), i.e. a model of the differential \( \lambda \)-calculus. Remember that the (qualitative) Taylor expansion of an ordinary \( \lambda \)-term \( M \) is an inductively defined series \( \mathcal{T}(M) \) of differential \( \lambda \)-terms, the only non-trivial case being \( \mathcal{T}(PQ) := \sum_{i=0}^{\infty} (D^i[P] \cdot Q^i)0 \). It can be seen that the morphisms of \( \mathbb{L}\text{Rel}_l \) can always be Taylor expanded, and the series interpreting in \( \mathbb{L}\text{Rel}_l \) the Taylor expansion of a STLC-term \( M \), converges to the interpretation of \( M \).
2 Tropical Laurent Series and Metric Semantics for STLC

The main goal of this section is to show that the interpretation of the above mentioned variants of the STLC in the tropical relational model yields a metric semantics, where the spaces $\mathbb{L}^X$ are endowed with the $\| \cdot \|_\infty$-norm metric, and programs are interpreted as locally Lipschitz maps:

**Theorem 1.** For any $\lambda$-term $M$:
1. if $\Gamma \vdash_{\text{STLC}} M : A$, then $[M] : [\Gamma] \rightarrow [A]$ is a Lipschitz map.
2. if $\Gamma \vdash_{\text{STLC}} M : A$, then $[M]^\dagger : [\Gamma] \rightarrow [A]$ is a locally Lipschitz map.

Moreover, the Taylor expansion $\mathcal{T}(M)$ decomposes $[M]^\dagger$ into an inf of Lipschitz maps.

Recall that the syntactic Taylor expansion decomposes an unbounded application as a limit of bounded ones; the result above lifts this decomposition to a semantic level, presenting a higher-order program as a decomposition of its components.

The proof of the result above requires the study of tLs from the point of view of mathematical analysis. We first studied tLs with respect to the topology induced by the $\| \cdot \|_\infty$-norm. This result shows that, in analogy with that happens in usual metric semantics, Lipschitz maps: $\mathcal{K}$ (in violet).

**Proposition 3.** If a tLs $f : \mathbb{L}^X \rightarrow \mathbb{L}^Y$ arises from a matrix $\hat{f} : \mathbb{L}^X \times \mathbb{L}^Y \rightarrow \mathbb{L}$, then $f$ is a $n$-Lipschitz map.

This result is perfectly analogous to what happens in the metric models recalled in the introduction. It also entails that any tropical polynomial $\varphi : \mathbb{L}^X \rightarrow \mathbb{L}$ is deg($\varphi$)-Lipschitz.

Let us now consider metric properties. First, it can be seen that all tropical linear functions $f : \mathbb{L}^X \rightarrow \mathbb{L}^Y$ are non-expansive. This result shows that, in analogy with that happens in usual metric semantics, linear programs are interpreted by non-expansive functions. More generally, linear maps with bounded exponentials yield Lipschitz maps:

**Proposition 4.** Let $k \in \mathbb{N}$ and $f : \mathbb{L}^k \rightarrow \mathbb{L}$ a tLs with matrix $\hat{f} : \mathbb{N}^k \rightarrow \mathbb{L}$. For all $0 < \varepsilon < \infty$, there is a finite $\mathcal{F}_\varepsilon \subseteq \mathbb{N}^k$ such that $f$ coincides on all $[\varepsilon, \infty]^k$ with the tropical polynomial $P_\varepsilon(x) := \inf_{n \in \mathcal{F}_\varepsilon} \{nx + \hat{f}(n)\}$.

The result above suggests that, far from 0, the behavior of tLs with finitely many variables can be studied with the tools of tropical geometry (e.g. tropical roots, Newton polygones). Moreover, a consequence of Theorem 4 is that all tLs with finitely many variables are always locally Lipschitz on $\mathbb{R}_{>0}$. The following result extends this property to all tLs, including the case with infinitely many variables.

**Theorem 5.** All tLs $\mathbb{L}^X \rightarrow \mathbb{L}$ are locally Lipschitz on $\mathbb{R}_{>0}^X$.  

Figure 1: Plot of the tropical polynomials $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ (from top to bottom), and of their limit tLs $\varphi$ (in violet).
The core of the proof is a simple convex analysis argument showing that an arbitrary function \( f : \mathbb{L}^X \to \mathbb{L} \) which is non-decreasing, concave and continuous, must be locally Lipschitz.

Finally, let us look at the differential operator \( D \) of \( \mathbb{L} \text{Rel} \). It translates into a differential operator \( D_t \) turing a \( \text{tL} f : \mathbb{L}^X \to \mathbb{L}^Y \) into a \( \text{tL} D_t f : \mathbb{L}^X \times \mathbb{L}^X \to \mathbb{L}^Y \), linear in its first variable, and given by \( D_t f(x,y)_b = \inf_{a \in X, \mu \in X} \{ f_{\mu + a} + x_a + \mu y \} \). One can check that, when \( f \) is a tropical polynomial, \( D_t f \) gives the standard tropical derivative (see e.g. \[16\]). Finally, the Taylor formula enjoyed by \( \mathbb{L} \text{Rel} \) morphisms, yields a “tropical” Taylor formula for \( \text{tL} \)s of the form \( f(x) = \inf_n \{ D_t^{(n)}(f)(1,x,\infty) \} \).

### 3 Tropical Semantics and Quantitative Properties: Likelihoods of Reduction Paths

Since algebraic and geometric properties in tropical mathematics are usually more tractable from a computational point of view, in several well-known applications (e.g. for optimization problems related to machine learning \[24, 28, 34\]) one starts from a given modelled phenomenon, typically expressed by some polynomial function \( f \), and studies which of its properties can be deduced from the tropicalization \( tf \) of \( f \). This idea suggests several natural directions in which the tropical semantics of a higher-order programs could be used to deduce properties which can be expressed as an optimization problem. We sketch here only one example, regarding probabilities, but we are currently working on other directions including differential privacy and best case analysis.

As a toy example, consider a probabilistic extension \( \text{STLC} \odot \) of \( \text{STLC} \), with a new ground type \( \text{Bool} \), terms \text{True}, \text{False} \,\text{of \, type \, Bool, terms \, of \, shape \,} M \odot_M N \,\text{and \, p}M, \,\text{for} \, p \in \{0, 1\}, \text{typed via the usual rules. We add reduction rules:} M \odot_M N \to pM \,\text{and} M \odot_M N \to (1-p)N, \,\text{so that} \, M \odot_M N \,\text{plays the role of} \, p. \,\text{STLC} \odot \,\text{can be seen as a fragment of the PCF in} \[19\].

Consider now the following term \( M := (\text{True} \oplus_p \text{False}) \oplus_p ((\text{True} \oplus_p \text{False}) \oplus_p (\text{False} \oplus_p \text{True})) \) of type \( \text{Bool} \). Let us give addresses \( \omega \in \{00, 01, 100, 101, 110, 111\} \) to the occurrences of \text{True}, \text{False} \,\text{in} \, M, \,\text{reading} \, 0 \,\text{as “left” and} \, 1 \,\text{as “right” in the tree structure of} \, M. \,\text{Calling} \, q := 1 - p, \,\text{there are the following six normal terms reachable from} \, M: P_{00}(p,q)\text{True}, P_{100}(p,q)\text{True}, P_{111}(p,q)\text{True}, P_{01}(p,q)\text{False}, P_{101}(p,q)\text{False}, P_{110}(p,q)\text{False}, \,\text{where the} \, P \,\text{’s are the following monomials in} \, p, q: P_{00}(p,q) := p^2, P_{01}(p,q) := q, P_{101}(p,q) := p, P_{110}(p,q) := q^2, P_{111}(p,q) := q^2. \,\text{They correspond to the respective reduction path from} \, M \,\text{to the normal term of the given address.} \, P_{\omega} \,\text{is then the probability (as a function of} \, p, q \,\text{) of obtaining the respective occurrence} \, \text{True}_\omega \,\text{or} \, \text{False}_\omega,\,\text{i.e. the likelihood function of the reduction given by} \, \omega. \,\text{The polynomial} \, Q_1(p,q) := P_{00}(p,q) + P_{100}(p,q) + P_{111}(p,q) \,\text{gives instead the whole probability of obtaining} \, \text{True} \,\text{after all the tossings (similarly a} \, Q_0(p,q) \,\text{for} \, \text{False}. \,\text{This way,} \,\text{the probabilistic evaluation of} \, M \,\text{is presented as a hidden Markov model} \[5\], \,\text{a fundamental statistical model, and notably one to which tropical methods are generally applied} \[28\]. \,\text{A typical question in this case would be: knowing that} \, M \,\text{produced} \, \text{True, which is the choice of the parameters} \, p, q \,\text{that maximizes the probability that among the paths leading to} \, \text{True, the one taken was a fixed} \, \omega_0? \,\text{Answering it, amounts at solving an optimisation problem related to} \, P_{\omega}, Q_{\omega}, \,\text{which is more easily solved via the tropicalizations} \, tP_{\omega}, tQ_{\omega}. \,\text{We are looking for} \, p, q \in [0, 1] \,\text{s.t.} \, q = 1 - p \,\text{and} \, \max_{\omega \in \{00, 100, 111\}} \, P_{\omega}(x,y) = P_{\omega_0}(x,y). \,\text{This last condition is equivalent to ask that} \, \min_{\omega \in \{00, 100, 111\}} -\log P_{\omega}(x,y) = -\log P_{\omega_0}(x,y), \,\text{i.e.} \,(tQ_1)(-\log x, -\log y) = (tP_{\omega_0})(-\log x, -\log y).

**Remark 1.** By adapting \[19\,\text{Section IV, it can be seen that} \, \mathbb{L} \text{Rel} \,\text{is a model of} \, \text{STLC} \odot. \,\text{In particular, if we set} \,[\text{Bool}] := \{0, 1\}, \,\text{our running example} \, M \,\text{is interpreted as} \,[M] \in \mathbb{L}^{\{0,1\}} \,\text{giving the following tropicalised probabilities:} \,[M]_0 = (tQ_0)(p, 1 - p), \,[M]_1 = (tQ_1)(p, 1 - p).
For our $M$, we have $tQ_1(x,y) = \min\{2x,y+2x,3y\}$ and $tQ_0(x,y) = \min\{x+y,2y+x\}$. Studying $tQ_1$, we see that $tQ_1(x,y) = 3y$ if $y \leq \frac{1}{2}x$, and it coincides with $2x$ otherwise. Remembering that $3y = P_{111}(x,y)$, we can now solve our optimisation problem above for $\omega_0 = 111$: via the substitution $x := -\log p$, $y := -\log(1-p)$, it is equivalent to $-\log(1-p) \leq -\frac{1}{2}\log p$, i.e. $1 - p \geq p^\frac{1}{2}$. This means that, for $p \in [0,1]$ s.t. $1 - p \geq p^\frac{1}{2}$ (for example, $p = \frac{1}{4}$), the most likely occurrence of True to obtain, knowing that $M$ sampled True in its normal form, is True$_{111}$. Remembering that $2x = P_{10}(x,y)$, for the other values of $p$ (for example, $p = \frac{1}{2}$), the most likely True to be sampled is the occurrence True$_{00}$. We have thus answered our question. Also, the $p \in [0,1]$ s.t. $(p, 1-p)$ is a root of $tQ_1$ or of $tQ_0$, provide the values of the bias of $\oplus_p$ for which there are at least two different equiprobable paths from $M$ to its normal form. Yet, we do not have a full understanding of the role of roots in this setting.

4 Getting rid of bases: $\mathbb{L}$-modules as generalized metric spaces

As we have seen, tropical semantics provides a viewpoint in which metric and differential properties coexist. This approach can be made more abstract, thanks to a fundamental categorical correspondence between tropical linear algebra (i.e. the study of quantale modules over $L$) and the theory of Lawvere’s generalized metric spaces (see [14],[32]).

A quantale module over $\mathbb{L}$, shortly a $\mathbb{L}$-module, is a triple $(M, \leq, \ast)$ where $(M, \leq)$ is a sup-lattice, and $\ast : \mathbb{L} \times M \rightarrow M$ is a continuous (left-)action of $\mathbb{L}$ on it, where continuous means that $\ast$ commutes with both joins in $\mathbb{L}$ and in $M$. The most basic examples of $\mathbb{L}$-modules are given by the spaces $\mathbb{L}^X$, with order and action defined pointwise. We see them as modules given together with a fixed base, $X$. Theorem 6 says that one can actually give a semantic of high order programs without need to fix a base.

We already remarked that the tropical semiring coincides with the Lawvere quantale $\mathbb{L}$. In particular, Lawvere was the first to observe that a (possibly $\infty$) metric on a set $X$ is nothing but a “$\mathbb{L}$-valued square matrix” $d : X \times X \rightarrow \mathbb{L}$ satisfying axioms like e.g. the triangular law. Indeed, such distance matrices correspond to $\mathbb{L}$-enriched categories (in short, a $\mathbb{L}$-category) [17],[20],[33]. Many topological properties of metric spaces translate in this way into purely categorical ones. In particular, $\mathbb{L}$-enriched categories which are cocomplete in the (enriched) categorical sense (i.e. all weighted colimits exist) satisfy usual metric completeness properties (e.g. Cauchy-completeness or Isbell-completeness).

Now, any $\mathbb{L}$-module $M$ can be endowed with a metric $M(x,y) = \inf\{\varepsilon \mid \varepsilon \ast x \geq y\}$, yielding a cocomplete $\mathbb{L}$-enriched category and, conversely, any cocomplete $\mathbb{L}$-enriched category $X$ has a $\mathbb{L}$-module structure with order $x \leq y$ when $X(y,x) = 0$, and monoidal operation defined via a suitable weighted colimit. This induces an isomorphism between the SMCC $\mathbb{L}$Mod of $\mathbb{L}$-modules and their homomorphisms and the SMCC $\mathbb{L}$CCat of cocomplete $\mathbb{L}$-enriched categories and cocontinuous functors (notice that functoriality in $\mathbb{L}$CCat is just non-expansiveness).

At this point, it can be shown that the SMCC structure of $\mathbb{L}$Rel lifts in a natural way to that of the more general category $\mathbb{L}$Mod $\simeq \mathbb{L}$CCat; secondly, an exponential ! in $\mathbb{L}$Mod $\simeq \mathbb{L}$CCat, can be defined via a well-known recipe based on the construction of symmetric algebras [18],[19],[27]. Since ! makes $\mathbb{L}$CCat a Lafont category, one can either do a few calculations or apply general theorems from the literature (e.g. [21]) to obtain the following result, which lifts also the exponential and differential structure of $\mathbb{L}$Rel to this more general setting:

**Theorem 6.** $\mathbb{L}$Mod or, equivalently, $\mathbb{L}$CCat, can be equipped with a differential $D$ making it a CC$\partial$C.
References


