About the Power of Taylor Expansion

Davide Barbarossa
LIPN, Université Paris 13
barbarossa@lipn.univ-paris13.fr

Giulio Manzonetto
LIPN, Université Paris 13
manzonetto@lipn.univ-paris13.fr

We discuss the real power of Ehrhard and Regnier's commutation result between Böhm trees and normalized Taylor expansion. It turns out that classic results like the contextuality of Böhm trees, syntactic continuity and the Parallel Lines Lemma can be given "soft" proofs by exploiting the main properties enjoyed by the resource calculus, namely strong normalization, confluence and linearity.

1 Introduction

The untyped λ-calculus has been studied for more than 80 years as an abstract model of computation, as well as a mathematical theory having its own interest [1]. Several fundamental results, such as the contextuality of Böhm trees, the Genericity Lemma and the Perpendicular Lines Lemma have been demonstrated by applying celebrated techniques, such as Scott's syntactic continuity, Berry's stability and Kahn and Plotkin's sequentiality theory, which are in principle non-trivial to establish.

In 2003 Ehrhard and Regnier performed a fine semantic analysis of Girard's linear logic [6] and observed that in Köthe spaces a derivative operator is actually at hand. This led them to introduce the *differential λ-calculus* [3] (and, subsequently, *differential proof-nets*) having a higher-order syntactic derivative operator that brings in the qualitative framework of λ-calculus a central idea of linear logic: resource sensitivity. In particular, it is possible to approximate the behaviour of a λ-term M by performing its *Taylor expansion*, an operation translating M into the powerseries[1] of all its linear approximants. The target language is a *resource calculus* similar to Boudol's λ-calculus with multiplicities, where a term is not applied to another term, but rather to a bag (finite multiset) of non-erasable nor dublicable terms [9]. A Commutation Theorem relates this way of approximating a λ-term with the more traditional one, based on Böhm trees approximants, domains and continuity: the theorem states that the normal form of the Taylor expansion of a λ-term coincides with the Taylor expansion of its Böhm tree [4].

The advantage of this approach is that – thanks to its solid background rooted on linear logic – it is easily generalizable to other settings, like algebraic, non-deterministic, probabilistic λ-calculi and programming paradigm like call-by-name, call-by-value, call-by-need, call-by-push value and this constitutes an ongoing axe of research followed by several people in the community. In the present work we show that this Commutation Theorem – which is non-trivial but not so difficult to demonstrate either – is actually very useful to establish the aforementioned results on Böhm trees and program behaviour, thus bypassing, or more precisely subsuming, the usual techniques as continuity, stability and sequentiality. In particular, we exploit the fact that the Taylor expansion has an inductive definition, as well as the main properties enjoyed by the resource calculus namely strong normalization, confluence and linearity in order to give "soft" proofs of these results. In particular, this allows to bring back the induction principle in the coinductive world of Böhm trees and it constitutes a convenient approach also when one it is not interested in studying quantitative properties of the programs, a widespread misconception. We believe that the same proof-techniques generalize to all frameworks and paradigms mentioned above.

[1] For our purposes we will be interested in the support of this expansion, i.e. the set of its summands

Submitted to:

2 Preliminaries

Concerning the λ-calculus we follow Barendregt's book [1]. So, the set Λ of *λ-terms* is defined by

$$\Lambda : \qquad M,N \ ::= \ x \mid \lambda x.M \mid MN$$

and endowed with β-reduction: $(\lambda x.M)N \to_\beta M\{N/x\}$ where the latter denotes capture free substitution. As usual, we denote by $\twoheadrightarrow_\beta$ and $=_\beta$ multistep β-reduction and β-conversion respectively and λ-terms are considered up to α-conversion. The *Böhm tree* $\mathrm{BT}(M)$ of a λ-term M is defined coinductively by:

$\mathrm{BT}(M) = \lambda x_1 \dots x_n.y\,\mathrm{BT}(M_1)\cdots \mathrm{BT}(M_k)$ if $M \twoheadrightarrow_\beta \lambda x_1 \dots x_n.yM_1 \cdots M_k$ for some $n,k \geq 0$.

$\mathrm{BT}(M) = \bot$ otherwise (i.e. when M is unsolvable, hence it does not have a head normal form);

The set App of *finite approximants* is defined inductively: the constant $\bot$ belongs to App; if $P_1,\dots,P_k \in \mathsf{App}$ then also $\lambda x_1 \dots x_n.P_1 \cdots P_k \in \mathsf{App}$. For $M \in \Lambda$ and $P \in \mathsf{App}$ we write $P \leq_\bot M$ if P is obtained from M by substituting some subterms with $\bot$. The set $\mathscr{A}(M)$ of *approximants of* M is defined by setting $\mathscr{A}(M) = \{P \in \mathsf{App} \mid M \twoheadrightarrow_\beta N \text{ and } P \leq_\bot N\}$. It is well known that $\mathscr{A}(M)$ is an ideal w.r.t. $\leq_\bot$ and $\mathrm{BT}(M) = \bigvee \mathscr{A}(M)$. We write $M =_\mathscr{B} N$ whenever $\mathscr{A}(M) = \mathscr{A}(N)$, equivalently, $\mathrm{BT}(M) = \mathrm{BT}(N)$.

We now recall the main notions, notations and properties concerning the resource calculus as well as the definition of Taylor expansion. The interested reader can read, e.g., [3] for a more detailed treatment.

Definition 1 (Resource calculus)**.** *The sets Λ^r of* resource terms *and Λ^b of* bags *are defined inductively:*

$$\Lambda^r : \ s,t \ ::= \ x \mid \lambda x.t \mid t\,[s_1,\dots,s_k] \qquad \textit{for } [s_1,\dots,s_k] \in \Lambda^b, \textit{ with } \Lambda^b := \mathscr{M}_f(\Lambda^r)$$

where $\mathscr{M}_f(X)$ denotes the set of finite multisets over X. We write $2\langle\Lambda^r\rangle$ for the set of finite formal sums of resource terms, quotiented by idempotency, symmetry and associativity of the sum. Thus, 0 denotes the empty sum and is the neutral element of the sum. As a syntactic sugar, we extend all the constructors to sums by (bi)linearity $(\sum_i t_i)(\sum_j s_j) := \sum_{ij} t_i s_i$ and $\lambda x.\sum_i t_i := \sum_i \lambda x.s_i$. In particular $\lambda x.0 = 0b = s0$.

Definition 2. *For $t \in \Lambda^r$ and $[s_1,\dots,s_k] \in \mathscr{M}_f(\Lambda^r)$, the* linear substitution *$t\langle [s_1,\dots,s_k]/x\rangle \in 2\langle\Lambda^r\rangle$ is:*

$$t\langle [s_1,\dots,s_k]/x\rangle := \begin{cases} \sum_{\sigma\in\mathfrak{S}_n} t\{s_{\sigma(1)}/x^{(1)},\dots,s_{\sigma(k)}/x^{(k)}\} & \textit{if } k = \deg_x(t) \\ 0 & \textit{otherwise} \end{cases}$$

where $\mathfrak{S}_n$ is the group of permutations over $\{1,\dots,n\}$, and $x^{(i)}$ we denote the i^{th} occurrence of x in t. The resource calculus is endowed with a reduction $\to_r \subseteq 2\langle\Lambda^r\rangle \times 2\langle\Lambda^r\rangle$ generated by $(\lambda x.t)b \to_r t\langle b/x\rangle$ and extended to sums by setting $t + \mathbb{T} \to_r t' + \mathbb{T}$ under the hypothesis that $t \to_r t'$.

It is well-known that such reduction is confluent and strongly normalizing, key properties that will be silently used in the following. So, the normal form $\mathrm{nf}(t) \in 2\langle\Lambda^r\rangle$ of $t \in \Lambda^r$ always exists (it can be 0).

Definition 3 (Taylor expansion)**.** *The* Taylor expansion *is the inductively defined map $\mathscr{T} : \Lambda \to \mathscr{P}(\Lambda^r)$:*

$$\mathscr{T}(x) := \{x\},\ \mathscr{T}(\lambda x.M) := \{\lambda x.t \mid t \in \mathscr{T}(M)\} \textit{ and } \mathscr{T}(MN) := \{tb \mid t \in \mathscr{T}(M) \text{ and } b \in \mathscr{M}_f(\mathscr{T}(N))\}.$$

This definition is extended to finite approximants setting $\mathscr{T}(\bot) := \emptyset$ and to Böhm trees setting:

$$\mathscr{T}(\mathrm{BT}(M)) := \bigcup_{P\in\mathscr{A}(M)} \mathscr{T}(P) \subseteq \Lambda^r$$

Also, for $M \in \Lambda$ we define the normal form of its Taylor expansion *as $\mathrm{NF}(\mathscr{T}(M)) := \sum_{t\in\mathscr{T}(M)} \mathrm{nf}(t)$, which we will usually identify with the set of its addends, namely a possibly infinite subset of Λ^r. For $M,N \in \Lambda$, we write $M \equiv N$ iff $\mathrm{NF}(\mathscr{T}(M)) = \mathrm{NF}(\mathscr{T}(N))$. One can show that $M =_\beta N$ implies $M \equiv N$.*

The following lemma will be used below and is straightforward to check by structural induction on P.

Lemma 1. *Let $M \in \Lambda$ and $P \in \mathsf{App}$. If $\mathscr{T}(P) \subseteq \mathscr{T}(\mathrm{BT}(M))$ then $P \in \mathscr{A}(M)$.*

We now state Ehrhard and Regnier's commutation theorem whose proof can be found in [4].

Theorem 1 (Commutation Theorem)**.** *For $M \in \Lambda$, we have $\mathrm{NF}(\mathscr{T}(M)) = \mathscr{T}(\mathrm{BT}(M))$.*

As an immediate corollary we obtain that M is solvable if and only if $\mathrm{NF}(\mathscr{T}(M)) \neq \emptyset$.

Corollary 1 (Characterization of the theory $\mathscr{B}$ of Böhm trees)**.** *Let $M,N \in \Lambda$. Then:*

$$\mathrm{BT}(M) \leq_\perp \mathrm{BT}(N) \iff \mathrm{NF}(\mathscr{T}(M)) \subseteq \mathrm{NF}(\mathscr{T}(N))$$

Proof. ($\Rightarrow$) This inclusion follows easily from Ehrhard and Regnier's Commutation Theorem.

($\Leftarrow$) Take $P \in \mathscr{A}(M)$ then $\mathscr{T}(P) \subseteq \mathscr{T}(\mathrm{BT}(M)) = \mathrm{NF}(\mathscr{T}(M)) \subseteq \mathrm{NF}(\mathscr{T}(N)) = \mathscr{T}(\mathrm{BT}(N))$, therefore by Lemma 1 and conclude $P \in \mathscr{A}(N)$. Thus $\mathscr{A}(M) \subseteq \mathscr{A}(N)$, equivalently $\mathrm{BT}(M) \leq_\perp \mathrm{BT}(N)$. □

This result immediately implies that $M =_{\mathscr{B}} N$ if and only if $M \equiv N$ and will be used in a crucial way.

3 Fundamental results proved via Taylor expansion normalization

We now show that it is possible to prove that the λ-theory $\mathscr{B} := \{(M,N) \mid \mathrm{BT}(M) = \mathrm{BT}(N)\} \subseteq \Lambda \times \Lambda$ of Böhm trees is contextual, passing through the monotonicity of the normal form of the Taylor expansion.

Proposition 1 (Monotonicity)**.** *Let $M,N \in \Lambda$.*

If $\mathrm{NF}(\mathscr{T}(M)) \subseteq \mathrm{NF}(\mathscr{T}(N))$ *then, for all contexts $C[]$, we have* $\mathrm{NF}(\mathscr{T}(C[M])) \subseteq \mathrm{NF}(\mathscr{T}(C[N]))$.

This immediately implies that if $M \equiv N$ then $C[M] \equiv C[N]$.

Proof. By structural induction on $C[]$. The only interesting case is $C[] = (C_1[])(C_2[])$. If $t \in \mathrm{NF}(\mathscr{T}(C[M]))$ then there exist $s \in \mathscr{T}(C_1[M])$ and $b \in \mathscr{M}_f(\mathscr{T}(C_2[M]))$ such that (by confluence) $sb \twoheadrightarrow_r \mathrm{nf}_\beta(s)\, \mathrm{nf}_\beta(b) \twoheadrightarrow_r t + \mathbb{T}$. The induction hypothesis gives $\mathrm{nf}(s) \subseteq \mathrm{NF}(\mathscr{T}(C_1[M])) \subseteq \mathrm{NF}(\mathscr{T}(C_1[N]))$ and, together with the monotonicity of the map $\mathscr{M}_f(\cdot)$, also $\mathrm{nf}(b) \subseteq \mathscr{M}_f(\mathrm{NF}(\mathscr{T}(C_2[M]))) \subseteq \mathscr{M}_f(\mathrm{NF}(\mathscr{T}(C_2[N])))$. Therefore, we conclude $t \in \mathrm{NF}(\mathscr{T}((C_1[N])(C_2[N])))$. □

Corollary 2 (Context Lemma for Böhm trees)**.** *Let $M,N \in \Lambda$.*

If $\mathrm{BT}(M) \leq_\perp \mathrm{BT}(N)$ *then, for all contexts $C[]$, we have* $\mathrm{BT}(C[M]) \leq_\perp \mathrm{BT}(C[N])$.

Proof. If $\mathrm{BT}(M) \subseteq \mathrm{BT}(N)$ then $\mathrm{NF}(\mathscr{T}(M)) \subseteq \mathrm{NF}(\mathscr{T}(N))$, therefore by monotonicity $\mathrm{NF}(\mathscr{T}(C[M])) \subseteq \mathrm{NF}(\mathscr{T}(C[N]))$ and so $\mathrm{BT}(C[M]) \leq_\perp \mathrm{BT}(C[N])$. □

The contextuality of $\mathscr{B}$ immediately follows, in other words $M =_{\mathscr{B}} N$ entails $C[M] =_{\mathscr{B}} C[N]$ for all contexts $C[]$. The following is a fundamental property of unsolvable λ-terms.

Proposition 2. *Let $C[]$ be a context. If U is unsolvable and $C[U]$ is solvable then $\forall M \in \Lambda$, $C[M]$ solvable.*

Proof. Since U is unsolvable we have $\mathrm{NF}(\mathscr{T}(U)) = \emptyset \subseteq \mathrm{NF}(\mathscr{T}(M))$, therefore, by monotonicity we derive $\mathrm{NF}(\mathscr{T}(C[M])) \supseteq \mathrm{NF}(\mathscr{T}(C[U])) \neq \emptyset$, whence we conclude that $C[M]$ is solvable. □

The meaning of the following theorem is that if N has a completely defined value and contains a meaningless subterm U, then U has no influence on the computation of this value, and can be replaced with any term. This motivates the identification "meaningless = unsolvable".

Theorem 2 (Genericity Lemma, [1, Prop. 14.3.24])**.** *Let $U \in \Lambda$ be unsolvable and $C[]$ be a context.*

$$\textit{If } C[U] \textit{ has a } \beta\textit{-normal form then } C[U] =_\beta C[M] \textit{ for all } M \in \Lambda.$$

Proof. $\mathrm{NF}(\mathscr{T}(U)) = \emptyset \subseteq \mathrm{NF}(\mathscr{T}(M))$, so $\mathrm{NF}(\mathscr{T}(C[U])) \subseteq \mathrm{NF}(\mathscr{T}(C[M]))$, so $\mathrm{BT}(C[U]) \subseteq \mathrm{BT}(C[M])$ and since $C[U]$ is β-normalizable it must be $\mathrm{BT}(C[U]) = \mathrm{BT}(C[M])$. □

The next theorem, called Parallel Lines Lemma, states that if a λ-term F seen as a function from Λ^n to Λ is constant (modulo some equality) on n perpendicular lines, then it must be constant everywhere. This statement was originally demonstrated for the λ-theory $\mathscr{B}$, and proved later to hold for β-conversion as well by Endrullis and de Vrijer [5] that applied van Daalen's *Reduction under Substitution* property [2], which is a strengthening of the famous "Barendregt Lemma". On the contrary, it was shown in [8] that such property fails in the closed term model of β-conversion due the existence of so-called Plotkin terms [7], namely β-distinct closed λ-terms M,N satisfying $ML =_\beta NL$ for all closed λ-term L.

We now sketch our proof of the original statement for Böhm trees, where the use of Taylor expansion allows to exploit induction and strong normalization despite the fact that we consider equality in $\mathscr{B}$.

Theorem 3 (Perpendicular Lines Lemma)**.** *Let $n \in \mathbb{N}$, $F, N_i, M_{ij} \in \Lambda$ for $i, j \in \{1,\dots,n\}$. If $\forall Z \in \Lambda$,*

$$\begin{cases} FM_{11}\cdots M_{1n-1}Z & =_{\mathscr{B}} \quad N_1 \\ FM_{21}\cdots Z \cdots M_{2n} & =_{\mathscr{B}} \quad N_2 \\ & \quad\vdots \\ FZM_{n2}\cdots M_{nn} & =_{\mathscr{B}} \quad N_n \end{cases} \tag{1}$$

then $\forall Z_1,\dots,Z_n \in \Lambda$,

$$FZ_1\cdots Z_n =_{\mathscr{B}} N_1 =_{\mathscr{B}} \cdots =_{\mathscr{B}} N_n$$

Idea of the Proof. Since we are considering equality in $\mathscr{B}$, the λ-term F when applied to a fresh variable z, as in Fz, can display a constant behaviour (as in the last equation) for several reasons:

1. the variable z is erased during the reduction as in $F = \lambda zy.y$;
2. the variable z is "hidden" behind an unsolvable as in $F = \lambda z.x(\Omega z)$;
3. the variable z is "pushed into infinity" as in $F = \mathsf{Y}(\lambda yzx.x(yz))$ where Y is a fixed point combinator. Such term satisfies $Fz =_\beta \lambda x.x(Fz)$ and for all $Fz \twoheadrightarrow_\beta F'$ we have that $z \in F'$ but $z \notin \mathrm{BT}(Fz)$.

Since $=_{\mathscr{B}}$ and $\equiv$ coincide, the hypotheses are that the λ-terms on the left and right hand-side of the equations above have the same normal form of the Taylor expansion, and it is enough to prove that the same holds for the λ-terms $FZ_1\cdots Z_n, N_1,\dots,N_n$ independently from the chosen $Z_1,\dots,Z_n \in \Lambda$. Working with resource approximants guarantees much stronger properties than those described for above. Indeed, while a λ-term Fz that is independent from z may exhibit all the aforementioned behaviours (1)-(3), a resource term $t[z,\dots,z]$ belonging to the Taylor expansion $\mathscr{T}(Fz)$ cannot erase nor push into infinity any occurrence of z in a reduction to a non-empty normal form, whence it can only be independent because the bag is actually empty. These are the intuitions behind the claim formulated below.

Claim 1. *For fresh variables $z_1,\dots,z_n$ and $tb_1\cdots b_n \in \mathscr{T}(F\vec{z})$, if $t\vec{b} \not\twoheadrightarrow_r 0$ then $b_1 = \cdots = b_n = []$.*

As the size of t strictly decreases along the reduction, this claim can be proved by induction on its size. By substitution $\{\vec{Z}/\vec{z}\}$ we immediately obtain $\mathrm{NF}(\mathscr{T}(FZ_1\cdots Z_n)) = \mathrm{NF}(\mathscr{T}(N_i))$ for all i. $\square$

From that, it follows immediately that the computation described in the λ-calculus is essentially sequential and not parallel. Actually, the same proof works with a minimal modification to show that the Perpendicular Lines Lemma holds also in the *closed* term model of $\mathscr{B}$, which seems to be an original result. With similar techniques, we are able to prove also some other important results (as usual, translating them from the language of Böhm trees to that of Taylor expansion, via the commutation formula).

Proposition 3 (Scott's continuity, [1, Thm. 14.4.10])**.** *Let $M, F \in \Lambda$. Then:*

$$\forall P \in \mathscr{A}(FM), \exists Q \in \mathscr{A}(M) \text{ s.t. } P \sqsubseteq_{\mathscr{B}} FQ$$

The meaning of the statement being that, when looking at a term F as a map from $\Lambda/_{=_{\mathscr{B}}}$ to itself, a finite portion (i.e. a $P \in \mathscr{A}(FM)$) of its output FM, can only be generated by a finite portion (i.e. a finite approximant $Q \in \mathscr{A}(M)$) of its input M.

Theorem 4 (Stability theorem [1, Thm. 14.4.10])**.** *Let $F, M_1, \ldots, M_n \in \Lambda$ and $\mathscr{X}_1, \ldots, \mathscr{X}_n \subseteq \Lambda$ s.t. for all $i \in \{1, \ldots, n\}$, $\exists L_i \in \Lambda$ s.t. $\forall N \in \mathscr{X}_i$, $\mathrm{BT}(N) \leq_\perp \mathrm{BT}(L_i)$.*

If $\forall i \in \{1, \ldots, n\}$, $\mathrm{BT}(M_i) = \bigcap_{N \in \mathscr{X}_i} \mathrm{BT}(N)$, then $\mathrm{BT}(FM_1\cdots M_n) = \bigcap_{\vec{N} \in \vec{\mathscr{X}}} \mathrm{BT}(FN_1\cdots N_n)$.

Here the intersection of Böhm trees is to be intended as the intersection of the set of approximants, and equalities are on the approximants as well.

References

[1] Henk P. Barendregt (1984): *The lambda-calculus, its syntax and semantics*, second edition. Studies in Logic and the Foundations of Mathematics 103, North Holland.

[2] Dirk van Daalen (1980): *The Language Theory of Automath*. Ph.D. thesis, Technical University Eindhoven.

[3] Thomas Ehrhard & Laurent Regnier (2003): *The differential lambda-calculus*. Theor. Comput. Sci. 309(1-3), pp. 1–41.

[4] Thomas Ehrhard & Laurent Regnier (2006): *Böhm Trees, Krivine's Machine and the Taylor Expansion of Lambda-Terms*. In: CiE, Lecture Notes in Computer Science 3988, Springer, pp. 186–197.

[5] Jörg Endrullis & Roel C. de Vrijer (2008): *Reduction Under Substitution*. In Andrei Voronkov, editor: Rewriting Techniques and Applications, 19th International Conference, RTA 2008, Lecture Notes in Computer Science 5117, Springer, pp. 425–440.

[6] Jean-Yves Girard (1987): *Linear Logic*. Theor. Comput. Sci. 50, pp. 1–102.

[7] Gordon D. Plotkin (1974): *The lambda-Calculus is ω-Incomplete*. Journal of Symbolic Logic 39(2), pp. 313–317.

[8] Richard Statman & Henk Barendregt (1999): *Applications of Plotkin-Terms: Partitions and Morphisms for Closed Terms*. J. Funct. Program. 9(5), pp. 565–575.

[9] Paolo Tranquilli (2009): *Nets Between Determinism and Nondeterminism*. Ph.D. thesis, Univ. Paris 7 and Univ. Roma 3.