Tracking Redexes in the lambda-calculus



jean-jacques.levy@inria.fr Irif 24 oct 2024

http://jeanjacqueslevy.net/talks/24track-rev/track.pdf



- for logicians: important tool for proof theory
- for computer scientists: kernel of functional programming

Tracking redexes in the lambda-calculus "The French School of Programming" ed. Bertrand Meyer, Springer, 2024.

- for logicians: important tool for proof theory
- for computer scientists: kernel of functional programming

RÉDUCTIONS CORRECTES ET OPTIMALES DANS LE LAMBDA-CALCUL		
Soutenue le 17 janvier 1978 devant la Commission composée de : MM L NOLIN Président		
	H BARENDREGT J Y GIRARD G HUET M NIVAT J C SIMON J VUILLEMIN	Examinateurs
	G KAHN	Invité

Tracking redexes in the lambda-calculus "The French School of Programming" ed. Bertrand Meyer, Springer, 2024.

Context-free languages

• ambiguous grammar

$$S \rightarrow S + S$$

 $S \rightarrow a$

• 2 distinct parse trees for a + a + a



• a parse tree represents a permutation equivalence class w.r.t. to derivations

 $S \rightarrow S + S \rightarrow a + S \rightarrow a + S + S \rightarrow a + a + S \rightarrow a + a + a$ $S \rightarrow S + S \rightarrow a + S \rightarrow a + S + S \rightarrow a + S + a \rightarrow a + a + a$ $S \rightarrow S + S \rightarrow S + S + S \rightarrow a + S + S \rightarrow a + a + S \rightarrow a + a + a$ $S \rightarrow S + S \rightarrow S + S + S \rightarrow S + a + S \rightarrow a + a + S \rightarrow a + a + a$

• only 1 leftmost derivation per parse tree

• calculus of functions with nested redexes and bound variables



• many permutations of redex contractions



local confluence



• let *R* and *S* be two occurences of redexes in a given term.

Then the **residuals** of *R* by *S* is the set R/S of occurences of (disjoint) redexes which remain of *R* when *S* is contracted.

• residuals of (underlined) redexes

$$\frac{(Ix)(Ix) \rightarrow (Ix)x}{\Delta (Ix) \rightarrow (Ix)(Ix)} \qquad Ix (Ix) \rightarrow x (Ix) \\ (\lambda x.Ix (Ix))y \rightarrow Iy (Iy) \\ \underline{\Delta (Ix)} \rightarrow (\Delta x) \qquad (Ix) \rightarrow x$$

where $\Delta = \lambda x.xx$, $I = \lambda x.x$

Equivalence by permutations

• 2 reductions are permutation equivalent by iterating the local confluence diagram



Equivalence by permutations

• example



$$\Delta = \lambda x.x x$$
$$F = \lambda f.f y$$
$$I = \lambda x.x$$
$$A = F I$$
$$B = I y$$



- single-redex reduction steps $M \xrightarrow{R} N$
- residuals S/R of another redex S in M are disjoint redexes
 - let ${\mathcal F}$ be a set of disjoint redexes
 - write $\rho : \mathcal{F}$ for any single-redex reduction ρ of redexes of \mathcal{F} in any order
 - these reductions are all cofinal (end on a same term)
- single-redex reductions are locally confluent



- moreover, let T be a another redex in M
- residuals of *T* on both sides of the permutation are the **same**



 $T/(R \sqcup S) = T/(S \sqcup R)$ the minicube lemma

T/(R;(S/R)) = T/(S;(R/S))

• definition with permutations

 \sim is the smallest equivalence relation such that:

(i)
$$R \sqcup S \sim S \sqcup R$$

(ii) $\rho \sim \sigma \implies \tau; \rho; v \sim \tau; \sigma; v$



aka parse trees for contex-free languages

- a standard reduction is an outside-in left-to-right reduction strategy
- any reduction is equivalent by permutations to a unique standard reduction



• standard reductions are canonical representatives in equivalence classes

Finite developments

• parallel reduction steps $M \xrightarrow{\mathcal{F}} N$ where \mathcal{F} is a set of redexes in a given term

finite development thm [Curry] let \mathcal{F} be a set of redexes in a term M

- all reductions contracting only residuals of \mathcal{F} have finite length
- these reductions are all cofinal (end on a same term *N*)
- these maximal reductions are named developments of ${\mathcal F}$
- write $\rho: M \xrightarrow{\mathcal{F}} N$ or simply $\rho: \mathcal{F}$ for any development ρ of \mathcal{F}
- residuals of any redex R in initial term are the same for all developments of \mathcal{F}

corollary: parallel moves [Curry]







Residual of reductions





parallel moves +

 $\rho \sqcup \sigma$ and $\sigma \sqcup \rho$ are cofinal

cube lemma + $\tau/(\rho \sqcup \sigma) = \tau/(\sigma \sqcup \rho)$

permutation equivalence (revisited) Let ρ and σ be coinitial reductions.

Then $\rho \sim \sigma$ iff $\rho/\sigma = \emptyset^m$ and $\sigma/\rho = \emptyset^n$



exercise Test previous example and counterexamples for permutation equivalence

Prefix modulo permutations

prefix modulo permutations Let ρ and σ be coinitial reductions.

Then $\rho \leq \sigma$ iff $\rho/\sigma = \emptyset^m$





exercise Prove $\rho \leq \sigma$ iff $\exists \tau, \ \rho; \tau \sim \sigma$

Prefix modulo permutations

lattice of prefix modulo permutations





Prefix modulo permutations

lattice of prefix modulo permutations



Easy lemmas

properties of permutations

(i)
$$\rho \sim \sigma \iff \forall \tau. \tau / \rho = \tau / \sigma$$

(ii) $\rho \sim \sigma \iff \rho / \tau \sim \sigma / \tau$
(iii) $\rho \sim \sigma \iff \tau; \rho \sim \tau; \sigma$
(iv) $\rho \sim \sigma \iff \rho; \tau \sim \sigma; \tau$
(v) $\rho \sqcup \sigma \sim \sigma \sqcup \rho$

proof

(i)
$$\rho \sim \sigma \implies \sigma/\rho = \emptyset^n \text{ and } \rho/\sigma = \emptyset^m$$

Thus $\tau/(\rho \sqcup \sigma) = \tau/(\rho; (\sigma/\rho)) = \tau/\rho/(\sigma/\rho) = \tau/\rho/\emptyset^n = \tau/O$
Similarly $\tau/(\sigma \sqcup \tau) = \tau/\sigma$
By cube lemma $\tau/\rho = \tau/\sigma$
Conversely, take $\tau = \rho$ and $\tau = \sigma$

Easy lemmas

properties of prefixes modulo permutations

(i)
$$\rho \leq \sigma \leq \rho \iff \rho \sim \sigma$$

(*ii*) $\rho \leq \sigma \leq \tau \implies \rho \leq \tau$

- (*iii*) $\rho \leq \rho \sqcup \sigma$ and $\sigma \leq \rho \sqcup \sigma$
- (*iv*) $\rho \leq \tau$ and $\sigma \leq \tau \implies \rho \sqcup \sigma \leq \tau$

(v)
$$\rho \leq \sigma \iff \tau; \rho \leq \tau; \sigma$$

proof

 $\rho \leq \tau \text{ and } \sigma \leq \tau \implies \rho \sqcup \sigma \leq \tau$



history: the terminology "equivalence by permutations" is due to Gérard Berry]
[my initial definition was with residual of reductions]



Initial motivation

- what is call-by-need in the λ -calculus ?
 - recursive program schemes with dag implementation [Vuillemin 1973]
 - attempts (non optimal) for the λ -calculus [Wadsworth 1972]
 - dag with 1st-order term-rewriting systems [Maranget 1992]
 - using explicit substitutions ? [Ariola et al 1995]
- problem with sharing of functions
- no single-redex reduction strategy can be optimal example $F_m(\lambda y. \Delta_n(yz))$ where $F_m = \lambda x. x I x x ... x$ and $\Delta_n = \lambda x. x x ... x$

Initial motivation

• no single-redex reduction strategy can be optimal



Initial motivation

- call-by-need is call-by-name with sharing to avoid duplications
- what is duplication ?
- duplication along reductions already performed
- so duplication should take care of history of reductions

Duplication of redexes

• call-by-need is call-by-name with sharing to avoid duplications



Duplication of redexes

• sharing of λ -abstractions ? $I = \lambda x. x$



• possible sharing !!



- sharing of contexts ? with boxes and tags ??
- back to easier theory ...

Redex & history

h-redex (definition) $\langle \rho, R \rangle$ when *R* is a redex in final term of ρ

residual of h-redex $\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle$ if there is τ such that $S \in R/\tau$



Residuals modulo permutations

• properties of residuals of h-redexes

(*i*)
$$\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \iff \rho \leq \sigma \land S \in R/(\sigma/\rho)$$

- (*ii*) $\langle \rho, R \rangle \lesssim \langle \rho, R \rangle$
- (*iii*) $\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \lesssim \langle \rho, R \rangle \iff \rho \sim \sigma \land R = S$
- (*iv*) $\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \lesssim \langle \tau, T \rangle \implies \langle \rho, R \rangle \lesssim \langle \tau, T \rangle$

$$(v) \quad \langle \rho, R \rangle \lesssim \langle \tau, T \rangle \land \rho \leq \sigma \leq \tau \implies \exists ! S, \ \langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \lesssim \langle \tau, T \rangle \quad \blacktriangleleft$$

(vi)
$$\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \iff \langle \tau; \rho, R \rangle \lesssim \langle \tau; \sigma, S \rangle$$

interpolation

• residuals of h-redexes are consistent with permutation equivalence

Duplication complete reductions

• a reduction step $\xrightarrow{\mathcal{F}}$ is duplication complete if \mathcal{F} is a maximal set of h-redexes residuals of a single h-redex



Goal [optimality thm] leftmost-outermost d-complete reductions are optimal in their number of reduction steps

```
• but how to find the \langle \sigma, S \rangle h-redex ?
```

labeled calculus

• labels over alphabet $\mathcal{A} = \{a, b, c, \cdots\}$



• labeled λ -calculus

 $M, N, \dots ::= x \mid MN \mid \lambda x.M \mid M^{\alpha}$ $(\lambda x.M)^{\alpha}N \to M\{x := N^{\lfloor \alpha \rfloor}\}^{\lceil \alpha \rceil}$ $M^{\alpha}\{x := N\} = M\{x := N\}^{\alpha}$ $(M^{\alpha})^{\beta} = M^{\alpha\beta}$

- again an alphabet of atomic labels $\mathcal{A} = \{a, b, c, \dots\}$
- their labeled λ -calculus [Asperti-Laneve]

$$M, N, \dots ::= x \mid MN \mid \lambda x.M \mid a:M$$
$$(a_1:a_2:\cdots a_n:\lambda x.M)N \to a_1:a_2:\cdots a_n:M\{x := a_n:a_{n-1}:\cdots a_1:N\}$$
$$(a:M)\{x := N\} = a:M\{x := N\}$$

• correspondence with paths in initial term



• Let
$$\Delta = \lambda x.xx$$
, $\gamma_1 = \lfloor a \rfloor$, $\gamma_2 = \gamma_1 \lfloor \gamma_1 \rfloor$
 $\Delta^a \Delta \to (\Delta^{\gamma_1} \Delta^{\gamma_1})^{\lceil a \rceil} \to (\Delta^{\gamma_2} \Delta^{\gamma_2})^{\lceil \gamma_1 \rceil \lceil a \rceil} \to \cdots$

• the name of a redex be the label of its function part

name($(\lambda x.M)^{\alpha}N$) = α

- the name of a redex gives its **origin**
- residuals of a redex keep their names
- created new redexes strictly contain the names of their creators
• An example:
$$\underline{\Delta} = \lambda x.(x^c x^d)^b, \ \Delta = \lambda x.(x^g x^h)^f$$

$$\begin{split} \Omega &= \underline{\Delta}^{a} \Delta^{e} \\ &\to \Omega_{1} = (\Delta^{\gamma_{1}} \Delta^{\delta_{1}})^{b\lceil a\rceil} & \gamma_{1} = e\lfloor a \rfloor c & \delta_{1} = e\lfloor a \rfloor d \\ &\to \Omega_{2} = (\Delta^{\gamma_{2}} \Delta^{\delta_{2}})^{f\lceil \gamma_{1}\rceil b\lceil a\rceil} & \gamma_{2} = \delta_{1}\lfloor \gamma_{1}\rfloor g & \delta_{2} = \delta_{1}\lfloor \gamma_{1}\rfloor h \\ &\to \Omega_{3} = (\Delta^{\gamma_{3}} \Delta^{\delta_{3}})^{f\lceil \gamma_{2}\rceil f\lceil \gamma_{1}\rceil b\lceil a\rceil} & \gamma_{3} = \delta_{2}\lfloor \gamma_{2}\rfloor g & \delta_{3} = \delta_{2}\lfloor \gamma_{2}\rfloor h \\ &\to \cdots \end{split}$$

• or simpler with partial labels: $\Delta = \lambda x.x x$

$$\begin{split} \Omega &= \Delta^{a} \Delta \\ \to \Omega_{1} &= (\Delta^{\gamma_{1}} \Delta^{\gamma_{1}})^{\lceil a \rceil} & \gamma_{1} &= \lfloor a \rfloor \\ \to \Omega_{2} &= (\Delta^{\gamma_{2}} \Delta^{\gamma_{2}})^{\lceil \gamma_{1} \rceil \lceil a \rceil} & \gamma_{2} &= \gamma_{1} \lfloor \gamma_{1} \rfloor \\ \to \Omega_{3} &= (\Delta^{\gamma_{3}} \Delta^{\gamma_{3}})^{\lceil \gamma_{2} \rceil \lceil \gamma_{1} \rceil \lceil a \rceil} & \gamma_{3} &= \gamma_{2} \lfloor \gamma_{2} \rfloor \\ \to \cdots \end{split}$$

- the labeled calculus is confluent
- the labeled calculus is strongly normalizable when reduction is restricted to
- a finite set of redex names
- unique normal form when exists
 - [Generalized Finite Developments thm]
- the standard λ -calculus can be seen as an infinite limit of finite labeled-calculi



confluence

a











- a standard reduction is an outside-in left-to-right reduction strategy
- any reduction can be reordered in a standard reduction [Curry 1958]



 λ -calculus



labeled λ -calculus

Permutation equivalence

• \sim corresponds to the coinitial / cofinal reductions of the labeled λ -calculus



Prefix modulo permutations

• \leq corresponds to reductions of the labeled λ -calculus



redex

families

Residuals modulo permutations

• residuals of h-redexes correspond to names of redexes in :





$$\begin{split} \Delta &= \lambda x. (x^{c} x^{d})^{b} \qquad A = (F^{i} I^{u})^{q} \\ F &= \lambda f. (f^{k} y^{\ell})^{j} \qquad B = (I^{\gamma} y^{\ell})^{q} \\ I &= \lambda x. x^{\nu} \qquad C = y^{\ell \lfloor \gamma \rfloor \nu \lceil \gamma \rceil q} \end{split}$$

family relation \simeq between h-redexes is defined by :

(*i*)
$$\langle \rho, R \rangle \lesssim \langle \sigma, S \rangle \implies \langle \rho, R \rangle \simeq \langle \sigma, S \rangle \simeq \langle \rho, R \rangle$$

(*ii*) $\langle \rho, R \rangle \simeq \langle \sigma, S \rangle \simeq \langle \tau, T \rangle \implies \langle \rho, R \rangle \simeq \langle \tau, T \rangle$



• symmetric + transitive closure of residuals modulo permutations

- from now on, we only consider standard reductions
- then the extraction relation on h-redexes is defined as follows

(*i*) $\langle o, R \rangle \triangleleft \langle o, R \rangle$ (*ii*) $\langle \rho, R \rangle \triangleleft \langle \sigma, S \rangle \implies \langle \rho', R' \rangle \triangleleft \langle H; \sigma, S \rangle$

where $\langle \rho', R' \rangle$ is defined by cases analysis on ρ w.r.t. *H*

• Case 1: ρ is in body of *H* or disjoint to the right of the contractum of *H* then ρ' is isomorphic to ρ



• Case 2: ρ is internal to an instance of a copy of the argument of *H* then ρ' is isomorphic to ρ in the argument of *H*



• Otherwise (*H* necessary for *R*)

$$ho' = H;
ho \wedge R' = R$$





• the family relation \simeq can be decided by extraction

 $\langle \rho, R \rangle \simeq \langle \sigma, S \rangle \iff \langle \tau, T \rangle \triangleleft \langle \rho, R \rangle \land \langle \tau, T \rangle \triangleleft \langle \sigma, S \rangle \text{ for some } \langle \tau, T \rangle$

- in fact $\langle \tau, T \rangle$ is unique and is the canonical representative of its family
- $\langle \tau, T \rangle$ is unique in family with minimum length of (standard) reduction τ

redexes are stable in the λ -calculus

• when $\langle \tau, T \rangle$ is a canonical representative

 $\langle \rho, R \rangle \simeq \langle \tau, T \rangle \land \tau \leq \rho \implies \langle \tau, T \rangle \lesssim \langle \rho, R \rangle$

- family complete reductions are duplication complete reductions
- sublattice of family complete reductions

[optimality thm] leftmost-outermost d-complete reductions are optimal in their number of reduction steps

• the family relation \simeq corresponds to names in the labeled calculus



• family complete reductions are labeled complete reductions

The sublattice of optimal reductions

• optimal family complete reductions



Generalized finite developments

GFD++ thm [JJL] let \mathcal{F} be a finite set of redex families

- all reductions contracting only h-redexes in \mathcal{F} have finite length
- these maximal reductions are permutation equivalent

corollary a lambda-term is strongly normalizable iff it only contracts a finite set of redex families.

strong normalization *finiteness* of redex families

Extra properties

- algebraic laws with parallel reductions of redexes
- residuals of parallel reductions
- optimality of family complete reductions
- reductions with ultra sharing [Lamping]
- linear logic without boxes [Gonthier et al]
- generalization to other systems (interaction systems, ...)

Conclusion

- real implementations of sharing (more than call-by-need) ?
 [non exponential implementations]
- subsets where possible manageable sharing (weak calculi, others?)
- intuitive proofs of strong normalization ($\lfloor \alpha \rightarrow \beta \rfloor = \alpha$, $\lceil \alpha \rightarrow \beta \rceil = \beta$, $\alpha\beta = \alpha$)
- simplification of the extraction process
- history-based information flow
- incremental computations (makefiles [Vesta], neural networks)