Reduction Strategies in the Lambda Calculus A Systematic Approach to Their Specification and Efficient Implementation with Abstract Machines

Małgorzata Biernacka

Institute of Computer Science, University of Wrocław

WEPN 2023

▶ one-step strategy is a function $F : \Lambda \to \Lambda$ s.t.

$$t o_{eta}^1 F(t)$$

examples: leftmost-outermost, call by name, call by value

- of interest both in theory and in practice of lambda calculus
- multitude of strategies, defined and studied in different disguises
- relevant for efficient computation in lambda calculus
- often used as folklore, auxiliary tool

- ► formats to define strategies
- methodology to interderive various semantics (based on functional programming)
- ► formalization and classification

Weak strategies

$$(\lambda xy.xy)((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{CbN}} \lambda y.(\lambda z.z)(\lambda w.w)y$$

Weak strategies

$$\frac{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))}{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))} \xrightarrow{\text{CbN}} \lambda y.(\lambda z.z)(\lambda w.w)y$$
$$\xrightarrow{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))} \xrightarrow{\text{CbV}} (\lambda xy.xy)(\lambda w.w)$$

Weak strategies

$$\frac{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))}{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))} \xrightarrow{\text{CbN}} \lambda y.(\lambda z.z)(\lambda w.w)y$$
$$\xrightarrow{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))} \xrightarrow{\text{CbV}} (\lambda xy.xy)(\lambda w.w) \xrightarrow{\text{CbV}} (\lambda y.(\lambda w.w)y)$$

- fully normalize terms
- \blacktriangleright need to descend under λ and account for free variables in terms
- conservative extensions of weak strategies (e.g., strong CbN, strong CbV, strong CbNeed)
- efficient implementations required e.g. for typechecking in dependent types

Normal order = "iterate CbN"

$$(\lambda xy.xy)((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{NO}} \lambda y.(\lambda z.z)(\lambda w.w)y$$

Normal order = "iterate CbN"

$$\frac{(\lambda xy.xy)((\lambda z.z)(\lambda w.w))}{\overset{\text{NO}}{\longrightarrow} \lambda y.(\lambda z.z)(\lambda w.w)y}$$
$$\overset{\text{NO}}{\longrightarrow} \lambda y.(\lambda w.w)y$$
$$\overset{\text{NO}}{\longrightarrow} \lambda y.y$$

 $(\lambda xy.x(yy))((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{SCbV}} (\lambda xy.x(yy))(\lambda w.w)$

Strong strategies

 $(\lambda xy.x(yy))((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{SCbV}} (\lambda xy.x(yy))(\lambda w.w)$ $\xrightarrow{\text{SCbV}} \lambda y. (\lambda w. w) (yy)$

$$\begin{array}{l} (\lambda xy.x(yy))((\lambda z.z)(\lambda w.w)) & \stackrel{\mathrm{SCbV}}{\longrightarrow} (\lambda xy.x(yy))(\lambda w.w) \\ & \stackrel{\mathrm{SCbV}}{\longrightarrow} \lambda y.(\lambda w.w)(yy) \\ & \stackrel{\mathrm{SCbV}}{\longrightarrow} \lambda y.yy \end{array}$$

$$(\lambda xy.x(yy))((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{SCbV}} (\lambda xy.x(yy))(\lambda w.w)$$
$$\xrightarrow{\text{SCbV}} \lambda y.(\lambda w.w)(yy)$$
$$\xrightarrow{\text{SCbV}} \lambda y.yy$$

Strong CbV = iterate Open CbV

$$(\lambda xy.x(yy))((\lambda z.z)(\lambda w.w)) \xrightarrow{\text{SCbV}} (\lambda xy.x(yy))(\lambda w.w)$$
$$\xrightarrow{\text{SCbV}} \lambda y.(\lambda w.w)(yy)$$
$$\xrightarrow{\text{SCbV}} \lambda y.yy$$

Strong CbV = iterate Open CbV

Strong CbV is nondeterministic – we can choose right-to-left normalization (rrSCbV)

Formats of operational semantics

- structural operational semantics
- big-step semantics
- reduction semantics
- abstract machine
- definitional interpreter

Big-step semantics

$$\frac{1}{\lambda x.t \Downarrow \lambda x.t} \quad \frac{t_1 \Downarrow \lambda x.t \quad t_2 \Downarrow t'_2 \quad t[x := t'_2] \Downarrow t'}{t_1 t_2 \Downarrow t'}$$

$$\frac{t_1 \Downarrow t'_1 \neq \lambda x.t \quad t_2 \Downarrow t'_2}{t_1 t_2 \Downarrow t'_1 t'_2}$$

Open call by value

Reduction semantics

$$w ::= \lambda x.t \mid x \, \vec{w} \qquad E ::= \Box \mid w \, E \mid E \, t$$
$$\frac{t \rightharpoonup_{\beta_w} t'}{(\lambda x.t) \, w \rightharpoonup_{\beta_w} t[x := w]} \qquad \frac{t \rightharpoonup_{\beta_w} t'}{E[t] \stackrel{lcbw}{\rightarrow} E[t']}$$

Left-to-right open call by value

- micro-step semantics (explicit decomposition and substitution)
- abstract model of language implementation
- work on source terms (not on compiled terms)
- constant cost of each transition
- abstract cost model of computation

Krivine machine for CbN evaluation

$$t ::= n \mid tt \mid \lambda t \qquad C ::= [t, E]$$

$$E ::= \bullet \mid C :: E$$

$$S ::= \bullet \mid C :: S$$

$$t \mapsto \langle t, \bullet, \bullet \rangle$$

$$\langle t_1 t_2, E, S \rangle \rightarrow \langle t_1, E, [t_2, E] :: S \rangle$$

$$\langle \lambda t, E, C :: S \rangle \rightarrow \langle t, C :: E, S \rangle$$

$$\langle 0, [t, E] :: E', S \rangle \rightarrow \langle t, E, S \rangle$$

$$\langle n+1, C :: E, S \rangle \rightarrow \langle n, E, S \rangle$$

- refocusing from reduction semantics to abstract machine
- functional correspondence from higher-order normalizer to abstract machine
- introduced for weak strategies, extendable to strong ones

Generalized reduction semantics

Normal order in λ -calculus

$$t ::= x \mid \lambda x. t \mid t t$$
 $a ::= x \mid a n$ $n ::= a \mid \lambda x. n$

$$\underline{E} ::= F \mid \lambda x.E \mid aE$$
$$F ::= \Box_F \mid F t$$

$$E[(\lambda x.t)s] \xrightarrow{no} E[t[x := s]]$$

Generalized reduction semantics - formalization

syntactic categories: kinds, initial kind. terms values, potential redices, elementary contexts parameterized by kinds **atomic plug** – defining meaning of contexts contraction function proofs of basic properties

```
Parameters (ckind term : Set)
       (init_ckind : ckind)
       (redex value : ckind -> Set).
Parameters
       (elem_ctx : ckind -> ckind -> Set)
       (elem_plug : \forall {k0 k1}, term ->
        elem ctx k0 k1 \rightarrow term).
Parameter contract :
       \forall {k}, redex k -> option term.
Axioms
(v_triv: \forall ..., ec: [t] = v -> \exists v', t = v')
(v_red: \forall \{k\} (v : value k) (r : redex k),
v <> r).
```

- generalized reduction semantics
- linear strict order <_{k,t} on instances of productions from P^k that are compatible with t (i.e., elementary k-contexts matching t)
- atomic decomposition functions
- conditions on input enforce unique decomposition

Input to generalized refocusing — normal order

elementary contexts

$$E ::= \lambda x . \Box_E \mid a \Box_E \mid \Box_F t$$
$$F ::= \Box_F t$$
search order $a \Box_E <_{E, _} \Box_F t$

Abstract machine for normal order

$$\begin{array}{rcl} \langle \lambda x.t, C, F \rangle_{\mathsf{e}} & \rhd & \langle C, F, \lambda x.t \rangle_{\mathsf{c}} \\ \langle \lambda x.t, C, E \rangle_{\mathsf{e}} & \rhd & \langle t, \lambda x.\Box :: C, E \rangle_{\mathsf{e}} \\ \langle t_1 t_2, C, k \rangle_{\mathsf{e}} & \rhd & \langle t_1, (k, \Box t_2) :: C, F \rangle_{\mathsf{e}} \\ \langle x, C, k \rangle_{\mathsf{e}} & \rhd & \langle C, k, x \rangle_{\mathsf{c}} \\ \langle \lambda x.\Box :: C, E, v \rangle_{\mathsf{c}} & \rhd & \langle C, E, \lambda x.v \rangle_{\mathsf{c}} \\ \langle n_e \Box :: C, E, v \rangle_{\mathsf{c}} & \rhd & \langle C, E, n_e v \rangle_{\mathsf{c}} \\ \langle (k, \Box s) :: C, F, \lambda x.t \rangle_{\mathsf{c}} & \rhd & \langle s, (k, x \Box) :: C, E \rangle_{\mathsf{e}} \end{array}$$

An abstract rewriting system $\langle S, \Rightarrow \rangle$ **traces** another system $\langle T, \rightarrow \rangle$ if there exists a surjection $[]: S \rightarrow T$ s.t.

- 1. if $s_1 \Rightarrow s_2$ then $\llbracket s_1 \rrbracket = \llbracket s_2 \rrbracket$ or $\llbracket s_1 \rrbracket \to \llbracket s_2 \rrbracket$
- 2. if $t_1 \rightarrow t_2$ then for each s_0 s.t. $\llbracket s_0 \rrbracket = t_1$ there exists
 - $s_0 \Rightarrow \ldots \Rightarrow s_{n+1}$, where $\llbracket s_0 \rrbracket = \ldots = \llbracket s_n \rrbracket$ and $\llbracket s_{n+1} \rrbracket = t_2$
- 3. there are no silent loops

Theorem

Let M – machine generated by generalized refocusing from a RS with terms T and reduction relation \rightarrow . Then M traces $\langle T, \rightarrow \rangle$.

Functional correspondence



```
type value = Value of (thunk -> value)
 and thunk = unit -> value
let rec eval (e : env) (t : term) : value =
  match t with
  | Var n -> List.nth e n ()
  | App(t0, t1) -> value_unfold (eval e t0)
                                 (fun () \rightarrow eval e t1)
               -> Value (fun v -> eval (v :: e) t)
  | Lam t
let main (t : term) : value = eval [] t
```

After closure conversion

```
type thunk = Thunk of term * env
and env = thunk list
type value = Closure of term * env
let rec eval (e : env) (t : term) : value =
 match t with
 | Var n -> let Thunk(t', e') = List.nth e n in
                  eval e' t'
  | App(t0, t1) -> let Closure(t', e') = eval e t0 in
                  eval (Thunk(t1, e)::e') t'
  Lam t -> Closure(t, e)
```

let main (t : term) : value = eval [] t

CPS translation

```
type thunk = Thunk of term * env
and env = thunk list
type value = Closure of term * env
let rec eval (e : env) (t : term) (k : value \rightarrow 'a) : 'a =
  match t with
  | Var n -> let Thunk(t', e') = List.nth e n in
                   eval e' t' k
  | App(t0, t1) -> eval e t0 (function Closure(t', e') ->
                   eval (Thunk(t1, e)::e') t' k)
               -> k (Closure(t, e))
  | Lam t
let main (t : term) : value = eval [] t (fun x \rightarrow x)
```

Defunctionalization of continuations

```
type thunk = Thunk of term * env
and env = thunk list
type value = Closure of term * env
type stack = thunk list
let rec eval (e : env) (t : term) (s : stack) : value =
  match t. s with
  | Var n, _ -> let Thunk(t', e') = List.nth e n in
                     eval e' t' s
  | Lam t, v :: s -> eval (v :: e) t s
  | App(t0, t1), _ -> eval e t0 (Thunk(t1, e) :: s)
  | Lam t, [] -> Closure(t, e)
```

let main (t : term) : value = eval [] t []

```
type thunk = Thunk of term * env
and env = thunk list
type conf = term * env * env
let transition (c:conf) : conf =
  match c with
  | (App(t0, t1), e, s) -> (t0, e, Thunk(t1, e) :: s)
  | (Lam t, e, v :: s) -> (t, v :: e, s)
  | (Var 0, Thunk(t, e) :: _, s) -> (t, e, s)
  | (Var n, _ :: e, s) -> (Var(n-1), e, s))
```

let load (t : term) : conf = (t, [], [])

- programming technique to produce full normal forms reduction-free, based on denotational semantics
- ▶ interpret terms in a model
- ▶ then reify semantic values into syntactic normal forms

Normalization by evaluation for CbV

```
type term = Var of index | Lam of term | App of term * term
type level = int
type sem = Abs of (sem -> sem) | Neutral of (level -> term)
let to sem (f : sem -> sem) : sem = Abs f
let from_sem (d : sem) : sem -> sem =
  fun d' ->
    match d with
    | Abs f ->
      f d'
    | Neutral 1 ->
      Neutral (fun m \rightarrow let n = reify d' m in App (1 m, n))
```

Normalization by evaluation for CbV

```
let rec eval (t : term) (e : sem list) : sem =
  match t with
  | Var n -> List.nth e n
  Lam t' -> to_sem (fun d -> eval t' (d :: e))
  | App (t1, t2) \rightarrow let d2 = eval t2 e
                      in from_sem (eval t1 e) d2
let rec reify (d : sem) (m : level) : term =
  match d with
  | Abs f ->
    Lam (reify (f (Neutral (fun m' \rightarrow Var (m'-m-1))))(m+1))
  | Neutral 1 ->
    l m
```

let nbe (t : term) : term = reify (eval t []) 0

Normalization by evaluation for CbV

- functional correspondence applied to CbV NbE produces AM performing full normalization in Strong CbV strategy
- from the machine we can read off the reduction contexts
- obtained AM is inefficient: does not reuse constructed structures and suffers from size explosion

Size explosion problem

$$\omega := \lambda x. x x$$
$$e_n := \lambda x. c_n \omega x$$

Under Strong CbV

 e_n normalizes in linear number of steps to normal form of exponential size

$$\begin{array}{cccc} \omega^1 \, x \ \rightarrow & x \, x \\ \omega^2 \, x \ \rightarrow^* & (x \, x) \, (x \, x) \\ \vdots \end{array}$$

```
type sem = Abs of (sem -> sem)
         | Neutral of (unit -> term)
         Cache of term cache * sem
let rec from sem : sem -> (sem -> sem) = function
  | Abs f
                          -> f
            Neutral l -> apply_neutral l
  | Cache (c, Neutral 1) -> apply_neutral
                              (fun () \rightarrow cached call c l)
  | Cache (c. v) -> from sem v
and apply_neutral (1 : unit -> term) (v : sem) : sem =
  Neutral (fun () \rightarrow let n = reify v in App (1 (), n))
```

```
let rec eval (t : term) (e : env) : sem =
  match t with
  | Var x -> env_lookup x e
  | Lam (x, t') -> to_sem
        (fun v -> eval t' @@ Dict.add x (mount_cache v) e)
  | App (t1, t2) \rightarrow let v2 = eval t2 e
                     in from sem (eval t1 e) v2
let rec reify : sem -> term = function
  Abs f -> let xm = "x_" ^ string_of_int (gensym ()) in
    Lam (xm, reify (f @@ abstract_variable xm))
  | Neutral l \rightarrow l ()
  Cache (c, v) -> cached_call c (fun () -> reify v)
```

Terms
$$t ::= x | t_1 t_2 | \lambda x. t$$

Values $v ::= V(x) | v_1 v_2 | [x, t, E] | v^{\ell}$
Frames $F ::= [t, E] \Box | \Box v | v \Box | \Box t | \lambda x. \Box | @[\ell]$
Heaps $H :$ location \rightarrow term option
Conf. $K ::= \langle t, E, S, m, H \rangle_{\mathcal{E}} | \langle S, v, m, H \rangle_{\mathcal{C}}$
 $| \langle S, t, m, H \rangle_{\mathcal{S}} | \langle t^?, \ell, S, v, m, H \rangle_{\mathcal{M}}$

RKNV – abstract machine for SCbV

RKNV traces right-to-left strong call by value
 normal forms are equal up to α-equivalence

- amortized cost analysis based on configuration potential
- potential Φ_K of configuration K = how many steps the machine can make till the next β-step
- ▶ all but one transitions decrease potential

Configuration potential

 $\Phi_{\mathsf{K}}(K) := \Phi_{\mathsf{t}}(t) + \Phi_{\mathsf{S}}(S) + \Phi_{\mathsf{H}}(K) \qquad \text{if } K = \langle t, E, S, m, H \rangle_{\mathcal{E}}$

Erasure transition (precedes β -step)

$$\langle \Box \ v :: S_1, [x, t, E]^{\ell}, m, H \rangle_{\mathcal{C}} \stackrel{\mathsf{pre}\beta}{\to} \langle \Box \ v :: S_1, [x, t, E], m, H \rangle_{\mathcal{C}}$$

reasonability Let: $|\rho|$ – number of transitions starting from term t_0 , n – the number of β -reductions in rrSCbV normalization of term t_0 . Then $|\rho| \le (n+1) \cdot \Phi_t(t_0)$. overall complexity $O((1+n) \cdot |t_0| \cdot E(|t_0|))$

E(n) – cost of operations on environment of size n

- Strong CbV can be simulated in polynomial time
- Strong CbV calculus is a reasonable time cost model using approach alternative to [Accattoli et al. '21]

Strong CbNeed – how to approach it

- extend weak CbNd
- ▶ two approaches for weak CbNd: storeless or store-based

- complex declarative definition of reduction semantics [Balabonski et al. '17]
- expressible with generalized reduction semantics with contexts parameterized with sets of variables [Biernacka et al. '19]
- AM for SCbNd derived from generalized reduction semantics by refocusing

Strong Call by Need – example

$$\lambda z. (\lambda x. x t) (z z)$$
 z is frozen
 $\rightarrow \lambda z. \text{let } x = z z \text{ in } x t$

$$\lambda z. (\lambda x. x t) (z z) \quad z \text{ is frozen}$$

$$\rightarrow \lambda z. \text{let } x = z z \text{ in } x t$$

$$\equiv \lambda z. \text{let } x = z z \text{ in } [x] t$$

$$\lambda z. (\lambda x. x t) (z z) z \text{ is frozen}$$

$$\rightarrow \lambda z. \text{ let } x = z z \text{ in } x t$$

$$\equiv \lambda z. \text{ let } x = z z \text{ in } [x] t$$

$$\rightarrow \lambda z. \text{ let } x := z z \text{ in } x t$$

$$\lambda z. (\lambda x. x t) (z z) z \text{ is frozen}$$

$$\rightarrow \lambda z. \text{let } x = z z \text{ in } x t$$

$$\equiv \lambda z. \text{let } x = z z \text{ in } [x] t$$

$$\rightarrow \lambda z. \text{let } x := z z \text{ in } x t$$

$$\equiv \lambda z. \text{let } x := z z \text{ in } [x] t x \text{ is frozen}$$

$$\lambda z. (\lambda x. x t) (z z) z \text{ is frozen}$$

$$\rightarrow \lambda z. \operatorname{let} x = z z \operatorname{in} x t$$

$$\equiv \lambda z. \operatorname{let} x = z z \operatorname{in} [x] t$$

$$\rightarrow \lambda z. \operatorname{let} x := z z \operatorname{in} x t$$

$$\equiv \lambda z. \operatorname{let} x := z z \operatorname{in} [x] t x \text{ is frozen}$$

$$\equiv \lambda z. \operatorname{let} x := z z \operatorname{in} x [t] \rightarrow \dots$$

Strong CbNeed – alternative approach

- starting point: NbE normalizer for normal order
- introduce memoized thunks to avoid recomputation of arguments and of normal forms
- ► apply functional correspondence to derive AM

Strong CbNeed – RKNL abstract machine

$$\begin{split} \langle [\lambda x. t, e], s, \sigma \rangle_{\nabla} & \to \langle [\lambda x. t, e]^{\ell}, s, \sigma * [\ell \mapsto \bot] \rangle_{\triangle} \\ \langle [\lambda x. t, e]^{\ell}, s, \sigma \rangle_{\triangle} & \to \\ & \langle [t, e[x := \ell_2]], \lambda \check{x}. \Box :: \mathfrak{O}[\ell] :: s, \sigma * [\ell_2 \mapsto \check{x}_{\checkmark}] \rangle_{\nabla} \\ & \text{ where } \sigma(\ell) = \bot \end{split}$$

$$\langle [\lambda x. t, e]^{\ell}, \Box [t_2, e_2] :: s, \sigma \rangle_{\vartriangle} \to \langle [t, e[x := \ell_2]], s, \sigma * [\ell_2 \mapsto [t_2, e_2]] \rangle_{\bigtriangledown}$$

- derived store-based AM
- RKNL simulates the normal-order strategy:
 - each machine configuration accounts for a sequence of NO reduction steps (modulo α-equivalence)
 - each NO reduction step is simulated by a sequence of machine steps
- amenable to complexity analysis using potential function: number of transitions bilinear in the number of β-steps in NO and in size of initial term

- collection of reduction strategies
- characterized by term decompositions
- based on ubiquitous reduction contexts
- ► formalized in Coq

Context

```
Inductive frame : Type :=

| Lam : string \rightarrow frame

| Rapp : term \rightarrow frame

| Lapp : term \rightarrow frame.
```

Definition context : Type := list frame.

```
Definition CBN_frame (f:frame) : Prop :=
match f with
| Rapp _ => True
=> False
end.
Fixpoint Uniform (F:frame \rightarrow Prop) (C:context) : Prop :=
match C with
| [] => True
| f :: C => F f \land Uniform F C
end.
```

Definition CBN : context \rightarrow **Prop** := Uniform CBN_frame.



```
Definition decomposition : Type := context * term.
Definition strategy : Type := decomposition \rightarrow Prop.
Definition recompose : decomposition \rightarrow term := uncurry plug.
Definition normal_form (s:strategy) (t:term) : Prop :=
\neg \exists d, t = recompose d \land d \in s.
```

```
Definition det_strategy (s:strategy) : Prop := \forall t, \exists \leq 1 d, t = recompose d \land d \in s.
```

```
Definition \beta_contrex : term \rightarrow Prop := app_of abstraction \bullet.
```

```
Definition cbn : strategy := CBN \times \beta_contrex.
```

Normal forms

```
Fixpoint rigid (t:term) : Prop :=
match t with
| var _ => True
| app s _ => rigid s
=> False
end.
Definition whnf : term \rightarrow Prop := abstraction \cup rigid.
Example rigid_is_whnf : rigid \subseteq whnf.
Lemma cbn_nf : normal_form cbn == whnf.
```

Zoo



```
Definition sequence_strategy (r s: strategy) : strategy := \lambda d, r d \vee (normal_form r (recompose d) \wedge s d).
```

```
Notation "\checkmark" := left_strategy.
Notation "'\beta'" := only_\beta_contraction.
```

```
Definition \operatorname{cbn_phased} := \swarrow \operatorname{cbn};; \beta.
```

Definition no_phased := (β ;; \swarrow no;; \searrow no) $\cup \downarrow$ no.

Definition $cbw_phased := (\swarrow cbw \cup \searrow cbw);; \beta$. Definition $scbw_phased := (cbw;; (\checkmark scbw \cup \searrow scbw)) \cup \downarrow scbw$. Lemma $scbw_conservative_extension_cbw : scbw == cbw;; scbw.$

- Lemma sequence_strategy_assoc : \forall q r s, q;; (r;; s) == (q;; r);; s.

Lemma phased_left_strategy : \forall s s', \swarrow s;; \checkmark s' == \checkmark (s;; s').

Benefits

- framework to study, compare and discover new strategies
- more structured and generic proofs of strategy properties, normal forms, etc.
- algebraic reasoning about strategies

Thank you!