A computational analysis of proof transformation by forcing

Alexandre Miquel

June 1st, 2010 – LAMA, Chambéry

Introduction

- The forcing technique :
 - Introduced by Cohen to prove $Cons(ZFC + \neg HC)$
 - Formulæ interpreted as sets of conditions (belonging to a fixed poset C)
 - Formula translation : $A \mapsto p \Vdash A$ $(p \in C)$
- Krivine's interpretation of forcing (in 2nd/3rd order arithmetic)
 - ullet Underlying program transformation $t\mapsto t^*$ (on Curry-style proof-terms)
 - Correctness expressed via generalized realizability structures
- The aims of this talk :
 - Rephrase the translation in $PA\omega^+$ (independently from realizability)
 - Present the underlying program transformation $t \mapsto t^*$ and study its computational contents
 - Reveal the underlying computation model (i.e. abstract machine)

Plan

- Introduction
- 2 Higher-order arithmetic (tuned)
- 3 The forcing transformation
- 4 The forcing machine
- Conclusion

The forcing transformation

- Introduction
- 2 Higher-order arithmetic (tuned)

Higher-order arithmetic (PA ω^+)

A multi-sorted language that allows to express

```
• Individuals (sort \iota)
• Propositions (sort \circ)
```

• Functions over individuals
$$(\iota \to \iota, \quad \iota \to \iota \to \iota, \quad ...)$$

• Predicates over individuals $(\iota \to o, \quad \iota \to \iota \to o, \quad ...)$

• Predicates over individuals
$$(\iota \to o, \iota \to \iota \to o, \ldots)$$

• Predicates over predicates... $((\iota \to o) \to o, \ldots)$

Syntax of sorts (kinds) and higher-order terms

Sorts
$$au, \sigma ::= \iota \mid o \mid \tau \to \sigma$$

Terms $M, N, A, B ::= x^{\tau} \mid \lambda x^{\tau} \cdot M \mid MN \mid 0 \mid s \mid rec_{\tau} \mid A \Rightarrow B \mid \forall x^{\tau} A \mid \langle M = M' \rangle A$

• Implication without computational contents : $\langle M = M' \rangle A$

ullet Provably equivalent to : $M=_{ au}M'\Rightarrow A$ (Leibniz equality)

Conversion (1/2)

- Conversion $M\cong_{\mathcal{E}} M'$ parameterized by a (finite) set of equations $\mathcal{E}\equiv M_1=M_1',\ldots,M_k=M_k'$ (non oriented, well sorted)
- Reflexivity, symmetry, transitivity + base case :

$$\overline{M\cong_{\mathcal{E}} M'}$$
 $(M=M')\in \mathcal{E}$

• β -conversion, recursion :

$$(\lambda x^{\tau} \cdot M)N \cong_{\mathcal{E}} M\{x := N\}$$

$$\operatorname{rec}_{\tau} M M' 0 \cong_{\mathcal{E}} M$$

$$\operatorname{rec}_{\tau} M M'(s N) \cong_{\mathcal{E}} M' N (\operatorname{rec}_{\tau} M M' N)$$

• Usual context rules + extended rule for $\langle M = M' \rangle A$:

$$\frac{A \cong_{\mathcal{E}, M = M'} A'}{\langle M = M' \rangle A \cong_{\mathcal{E}} \langle M = M' \rangle A'}$$

Rules for identifying (computationally equivalent) propositions :

 $\forall x^{\tau} \, \forall y^{\sigma} \, A \quad \cong_{\mathcal{E}} \quad \forall y^{\sigma} \, \forall x^{\tau} \, A$

$$\forall x^{\tau} A \cong_{\mathcal{E}} A \qquad x^{\tau} \notin FV(A)$$

$$A \Rightarrow \forall x^{\tau} B \cong_{\mathcal{E}} \forall x^{\tau} (A \Rightarrow B) \qquad x^{\tau} \notin FV(A)$$

$$\langle M = M' \rangle \langle N = N' \rangle A \cong_{\mathcal{E}} \langle N = N' \rangle \langle M = M' \rangle A$$

$$\langle M = M \rangle A \cong_{\mathcal{E}} A$$

$$A \Rightarrow \langle M = M' \rangle B \cong_{\mathcal{E}} \langle M = M' \rangle (A \Rightarrow B)$$

$$\forall x^{\tau} \langle M = M' \rangle A \cong_{\mathcal{E}} \langle M = M' \rangle \forall x^{\tau} A \qquad x^{\tau} \notin FV(M, M')$$

• Example : $\top := \langle \mathsf{tt} = \mathsf{ff} \rangle \bot$ (type of all proof-terms) where $tt \equiv \lambda x^{o} y^{o} \cdot x$, $ff \equiv \lambda x^{o} y^{o} \cdot y$ and $\bot \equiv \forall z^{o} z$

Deduction system (typing)

```
• Proof terms : t, u := x \mid \lambda x \cdot t \mid tu \mid \infty
                                                                          (Curry-style)
```

• Contexts:
$$\Gamma ::= x_1 : A_1, \dots, x_n : A_n$$
 (A_i of sort o)

Deduction/typing rules

$$\frac{\mathcal{E}; \Gamma \vdash \mathbf{x} : A}{\mathcal{E}; \Gamma \vdash \mathbf{x} : A} \stackrel{(\mathbf{x} : A) \in \Gamma}{\underbrace{\mathcal{E}; \Gamma \vdash \mathbf{t} : A'}} \frac{\mathcal{E}; \Gamma \vdash \mathbf{t} : A}{\mathcal{E}; \Gamma \vdash \mathbf{t} : A'} \stackrel{A \cong_{\mathcal{E}} A'}{\underbrace{\mathcal{E}; \Gamma \vdash \mathbf{t} : A'}}$$

$$\frac{\mathcal{E}; \Gamma, \mathbf{x} : A \vdash \mathbf{t} : B}{\mathcal{E}; \Gamma \vdash \lambda \mathbf{x} . \mathbf{t} : A \Rightarrow B} \frac{\mathcal{E}; \Gamma \vdash \mathbf{t} : A \Rightarrow B}{\mathcal{E}; \Gamma \vdash \mathbf{t} : B}$$

$$\frac{\mathcal{E}, M = M'; \Gamma \vdash \mathbf{t} : A}{\mathcal{E}; \Gamma \vdash \mathbf{t} : \langle M = M' \rangle A} \frac{\mathcal{E}; \Gamma \vdash \mathbf{t} : \langle M = M \rangle A}{\mathcal{E}; \Gamma \vdash \mathbf{t} : A}$$

$$\frac{\mathcal{E}; \Gamma \vdash \mathbf{t} : A}{\mathcal{E}; \Gamma \vdash \mathbf{t} : \forall \mathbf{x}^{\tau} A} \frac{\mathcal{E}; \Gamma \vdash \mathbf{t} : \forall \mathbf{x}^{\tau} A}{\mathcal{E}; \Gamma \vdash \mathbf{t} : A \{\mathbf{x} := N^{\tau}\}}$$

$$\frac{\mathcal{E}; \Gamma \vdash \mathbf{c} : ((A \Rightarrow B) \Rightarrow A) \Rightarrow A}{\mathcal{E}; \Gamma \vdash \mathbf{c} : (A \Rightarrow B) \Rightarrow A} \Rightarrow A$$

All proof-terms have type $\top \equiv \langle \mathsf{tt} = \mathsf{ff} \rangle \bot$ (normalization fails)

From operational semantics...

- Krivine's λ_c -calculus
 - λ -calculus with call/cc and continuation constants :

$$t, u ::= x \mid \lambda x . t \mid tu \mid \mathbf{c} \mid \mathbf{k}_{\pi}$$

The forcing transformation

An abstract machine with explicit stacks :

• Stack = list of closed terms (notation :
$$\pi$$
, π')

 \bullet Process = closed term \star stack

Evaluation rules

(weak head normalization, call by name)

(Grab)	$\lambda x \cdot t$	*	$u \cdot \pi$	\succ	$t\{x:=u\}$	*	π
(Push)	tu	*	π	\succ	t	*	$u \cdot \pi$
(Call/cc)	œ	*	$t\cdot\pi$	\succ	t	*	$k_\pi \cdot \pi$
(Resume)	k_{π}	*	$t\cdot\pi'$	\succ	t	*	π

... to classical realizability semantics

- Interpreting higher-order terms :
 - Individuals interpreted as natural numbers
 - Propositions interpreted as falsity values
 - Functions interpreted set-theoretically

$$\begin{bmatrix} \iota \end{bmatrix} = \mathbb{N} \\
 \llbracket o \end{bmatrix} = \mathfrak{P}(\Pi) \\
 \llbracket \tau \to \sigma \rrbracket = \llbracket \sigma \rrbracket^{\llbracket \tau \rrbracket}$$

• Parameterized by a pole $\perp \!\!\! \perp \subseteq \Lambda_c \star \Pi$

(closed under anti-evaluation)

Interpreting logical constructions :

The forcing transformation

Adequacy

lf

•
$$\mathcal{E}$$
; $x_1 : A_1, \dots, x_n : A_n \vdash t : B$ (in $PA\omega^+$)

$$\bullet \ \rho \models \mathcal{E}, \quad u_1 \in \llbracket A_1 \rrbracket_{\rho}^{\perp}, \ldots, u_n \in \llbracket A_n \rrbracket_{\rho}^{\perp}$$

then: $t\{x_1 := u_1; \dots; x_n := u_n\} \in [\![B]\!]_a^{\perp}$

Plan

- Introduction
- 2 Higher-order arithmetic (tuned)
- The forcing transformation
- 4 The forcing machine
- Conclusion

Representing conditions

- Intuition : Represent the set of conditions as an upwards closed subset of a meet-semilattice
- Take :
 - A sort κ of conditions, equipped with
 - A binary product $(p,q) \mapsto pq$ (of sort $\kappa \to \kappa \to \kappa$)
 - A unit 1 (of sort κ)
 - A predicate $p \mapsto C[p]$ of well-formedness (of sort $\kappa \to o$)
- Typical example : finite functions from τ to σ are modelled by
 - (binary relations $\subseteq \tau \times \sigma$) • $\kappa \equiv \tau \rightarrow \sigma \rightarrow o$
 - $pq \equiv \lambda x^{\tau} y^{\sigma} . p x y \vee q x y$ (union of relations p and q)
 - 1 $\equiv \lambda x^{\tau} v^{\sigma} . \bot$ (empty relation)
 - $C[p] \equiv "p$ is a finite function from τ to σ "

Combinators

- The forcing translation is parameterized by
 - \bullet The sort κ + closed terms $\cdot,$ 1, \emph{C}

• 9 closed proof terms $\alpha_*, \alpha_1, \ldots, \alpha_8$

(logical level) (computational level)

 $\alpha_* : C[1]$ $\alpha_1 : \forall p^{\kappa} \forall q^{\kappa} (C[pq] \Rightarrow C[p])$

 $\alpha_2 : \forall p^{\kappa} \forall q^{\kappa} (C[pq] \Rightarrow C[q])$

 $\alpha_3 : \forall p^{\kappa} \forall q^{\kappa} (C[pq] \Rightarrow C[qp])$

 $\alpha_4 : \forall p^{\kappa} (C[p] \Rightarrow C[pp])$

 $\alpha_5 : \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[(pq)r] \Rightarrow C[p(qr)])$ $\alpha_6 : \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[p(qr)] \Rightarrow C[(pq)r])$

 $\alpha_6 : \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[p(qr)] \Rightarrow C[(pq)r]$ $\alpha_7 : \forall p^{\kappa} (C[p] \Rightarrow C[p1])$

 α_8 : $\forall p$ $(C[p] \Rightarrow C[p])$ α_8 : $\forall p^{\kappa} (C[p] \Rightarrow C[1p])$

This set is not minimal. One can take α_* , α_1 , α_3 , α_4 , α_5 , α_7 and define : $\alpha_2 := \alpha_1 \circ \alpha_3$, $\alpha_6 := \alpha_3 \circ \alpha_5 \circ \alpha_3 \circ \alpha_5 \circ \alpha_3$, $\alpha_8 := \alpha_3 \circ \alpha_7$

Derived combinators

• The combinators $\alpha_1, \ldots, \alpha_8$ can be composed :

```
Example: \alpha_1 \circ \alpha_6 \circ \alpha_3: \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[(pq)r] \Rightarrow C[rp])
```

We will also use the following derived combinators :

```
: \forall p^{\kappa} \ \forall q^{\kappa} \ \forall r^{\kappa} \ (C[(pq)r] \Rightarrow C[pr])
                          \alpha_3 \circ \alpha_1 \circ \alpha_6 \circ \alpha_3
\alpha_9
                                                                                     : \forall p^{\kappa} \ \forall q^{\kappa} \ \forall r^{\kappa} \ (C[(pq)r] \Rightarrow C[qr])
\alpha10
                          \alpha_2 \circ \alpha_5
                                                                                     : \forall p^{\kappa} \ \forall q^{\kappa} \ (C[pq] \Rightarrow C[p(pq)])
\alpha_{11}
             := \alpha_9 \circ \alpha_4
                                                                                     : \forall p^{\kappa} \ \forall q^{\kappa} \ \forall r^{\kappa} \ (C[p(qr)] \Rightarrow C[q(rp)])
             := \alpha_5 \circ \alpha_3
\alpha_{12}
                                                                                  : \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[p(qr)] \Rightarrow C[(rp)q])
\alpha_{13}
             := \alpha_3 \circ \alpha_{12}
              := \alpha_5 \circ \alpha_3 \circ \alpha_{10} \circ \alpha_4 \circ \alpha_2 : \forall p^{\kappa} \forall q^{\kappa} \forall r^{\kappa} (C[p(qr)] \Rightarrow C[q(rr)])
\alpha_{14}
                                                                                     : \forall p^{\kappa} \ \forall a^{\kappa} \ \forall r^{\kappa} \ (C[p(ar)] \Rightarrow C[ap])
             :=
                          \alpha_0 \circ \alpha_3
\alpha_{15}
```

Important remark :

- $C[pq] \Rightarrow C[p] \land C[q]$, but $C[p] \land C[q] \not\Rightarrow C[pq]$ (in general)
- Two conditions p and q are compatible when C[pq]

• Let
$$p \le q := \forall r^{\kappa}(C[pr] \Rightarrow C[qr])$$

ullet \leq is a preorder with greatest element 1:

$$\begin{array}{lll} \lambda c \cdot c & : & \forall p^{\kappa} \ (p \leq p) \\ \lambda x y c \cdot y (x c) & : & \forall p^{\kappa} \ \forall q^{\kappa} \ \forall r^{\kappa} \ (p \leq q \Rightarrow q \leq r \Rightarrow p \leq r) \\ \alpha_8 \circ \alpha_2 & : & \forall p^{\kappa} \ (p \leq 1) \end{array}$$

• Product pq is the l.u.b. of p and q:

```
\begin{array}{cccc} \alpha_9 & : & \forall p^\kappa \ \forall q^\kappa \ (pq \leq p) \\ \alpha_{10} & : & \forall p^\kappa \ \forall q^\kappa \ (pq \leq q) \\ \lambda xy \, . \, \alpha_{13} \circ y \circ \alpha_{12} \circ x \circ \alpha_{11} & : & \forall p^\kappa \ \forall q^\kappa \ \forall r^\kappa \ (r \leq p \Rightarrow r \leq q \Rightarrow r \leq pq) \end{array}
```

• C (set of 'good' conditions) is upwards closed :

$$\lambda x c \cdot \alpha_1 (x (\alpha_7 c)) : \forall p^{\kappa} \forall q^{\kappa} (p \leq q \Rightarrow C[p] \Rightarrow C[q])$$

• Bad conditions are smallest elements :

$$\lambda x c \cdot x (\alpha_1 c) : \forall p^{\kappa} (\neg C[p] \Rightarrow \forall q^{\kappa} p \leq q)$$

The auxiliary translation $(_)^*$

• Translating sorts : $\tau \mapsto \tau^*$

$$\iota^* \equiv \iota$$
 $o^* \equiv \kappa \to o$ $(\tau \to \sigma)^* \equiv \tau^* \to \sigma^*$

Intuition: Propositions become sets of conditions

• Translating terms : $M \mapsto M^*$

Lemma

- $(M\{x^{\tau} := N\})^* \equiv M^*\{x^{\tau^*} := N^*\}$ (substitutivity)
- If $M_1 \cong_{\mathcal{E}} M_2$, then $M_1^* \cong_{\mathcal{E}^*} M_2^*$ (compatibility with conversion)

• Given a proposition A and a condition p, let:

$$p \Vdash A := \forall r^{\kappa}(C[pr] \Rightarrow A^*r)$$

The forcing translation is trivial on ∀ and ⟨_ = _⟩_ :

$$\begin{array}{ccc} \rho \Vdash \forall x^{\tau} A & \cong_{\varnothing} & \forall x^{\tau^*} (p \Vdash A) \\ \rho \Vdash \langle M_1 = M_2 \rangle A & \cong_{\varnothing} & \langle M_1^* = M_2^* \rangle (p \Vdash A) \end{array}$$

All the complexity lies in implication!

(cf next slide)

General properties

$$\beta_1 := \lambda xyc \cdot y (x c) : \forall p^{\kappa} \forall q^{\kappa} (q \leq p \Rightarrow (p \Vdash A) \Rightarrow (q \Vdash A))$$

$$\beta_2 := \lambda x c . x (\alpha_1 c) : \forall p^{\kappa} (\neg C[p] \Rightarrow p \Vdash A)$$

$$\beta_3 := \lambda x c \cdot x (\alpha_9 c) : \forall p^{\kappa} \forall q^{\kappa} ((p \Vdash A) \Rightarrow (pq \Vdash A))$$

$$\beta_4 := \lambda x c \cdot x (\alpha_{10} c) : \forall p^{\kappa} \forall q^{\kappa} ((q \Vdash A) \Rightarrow (pq \Vdash A))$$

• Definition of $p \Vdash A \Rightarrow B$ looks strange :

$$p \Vdash A \Rightarrow B \equiv \forall r^{\kappa}(C[pr] \Rightarrow (A \Rightarrow B)^{*}r)$$

$$\cong_{\varnothing} \forall r^{\kappa}(C[pr] \Rightarrow \forall q^{\kappa} \forall r'^{\kappa} \langle r = qr' \rangle ((q \Vdash A) \Rightarrow B^{*}r'))$$

But it is equivalent to

$$\forall q \, ((q \Vdash A) \Rightarrow (pq \Vdash B)) \qquad \left(\mathsf{Hint} : \begin{array}{c} p \Vdash A \Rightarrow B & q \Vdash A \\ \hline pq \Vdash B \end{array} \right)$$

Coercions between $p \Vdash A \Rightarrow B$ and $\forall q ((q \Vdash A) \Rightarrow (pq \Vdash B))$

$$\gamma_1 := \lambda x c y . x y (\alpha_6 c) : (\forall q ((q \Vdash A) \Rightarrow (pq \Vdash B)) \Rightarrow p \Vdash A \Rightarrow B)$$

$$\gamma_2 := \lambda xyc.x(\alpha_5 c)y : (p \Vdash A \Rightarrow B) \Rightarrow \forall q ((q \Vdash A) \Rightarrow (pq \Vdash B))$$

$$\gamma_3 := \lambda xyc.x(\alpha_{11}c)y : (p \Vdash A \Rightarrow B) \Rightarrow (p \Vdash A) \Rightarrow (p \Vdash B)$$

$$\gamma_4 := \lambda x c y \cdot x \left(y \left(\alpha_{15} c \right) \right) : \neg A^* p \Rightarrow p \Vdash A \Rightarrow B$$

Translating proof-terms

• Krivine's program transformation $t \mapsto t^*$:

- The translation inserts
 - γ_1 ("fold") in front of every abstraction
 - γ_3 ("apply") in front of every application
- A bound occurrence of x in t is translated as $\beta_3^n(\beta_4x)$, where *n* is the de Bruijn index of this occurrence

Soundness (in $PA\omega^+$)

 $\mathcal{E}: x_1: A_1, \ldots, x_n: A_n \vdash t: B$ then \mathcal{E}^* ; $x_1 : (p \Vdash A_1), \ldots, x_n : (p \Vdash A_n) \vdash t^* : (p \Vdash B)$

Computational meaning of the transformation

• A proof of $p \Vdash A \equiv \forall r^{\kappa}(C[pr] \Rightarrow A^*r)$ is a function waiting an argument c: C[pr] (for some $r) \rightsquigarrow$ computational condition

$$(\lambda x \cdot t)^* \quad \star \quad c \cdot u \cdot \pi \qquad \succ \qquad t^{\dagger} \{ x := u \} \quad \star \quad \alpha_6 \, c \cdot \pi$$

$$(tu)^* \quad \star \quad c \cdot \pi \qquad \succ \qquad \qquad t^* \quad \star \quad \alpha_{11} \, c \cdot u^* \cdot \pi$$

$$cc^* \quad \star \quad c \cdot t \cdot \pi \qquad \succ \qquad \qquad t \quad \star \quad \alpha_{14} \, c \cdot k_{\pi}^* \cdot \pi$$

$$k_{\pi}^* \quad \star \quad c \cdot t \cdot \pi' \qquad \succ \qquad \qquad t \quad \star \quad \alpha_{15} \, c \cdot \pi$$

where :

$$t^{\dagger} \equiv t^* \{ x := \beta_4 x \} \{ x_i := \beta_3 x_i \}_{i=1}^n$$

$$k_{\pi}^* \equiv \gamma_4 k_{\pi}$$

Evaluation combinators

```
lpha_6: C[p(qr)] \Rightarrow C[(pq)r]

lpha_{11}: C[pr] \Rightarrow C[p(pr)]

lpha_{14}: C[p(qr)] \Rightarrow C[q(rr)]

lpha_{15}: C[p(qr)] \Rightarrow C[qp]
```

Plan

- Introduction
- 2 Higher-order arithmetic (tuned)
- 4 The forcing machine

Real mode :

Forcing mode :

Adequacy in real and forcing modes

- New abstract machine means :
 - New classical realizability model (based on the KFAM)
 - New adequacy results

Adequacy (real mode)

lf

- $\mathcal{E}: x_1: A_1, \dots, x_n: A_n \vdash t: B$ (in $PA\omega^+$)
- \bullet $\rho \models \mathcal{E}, c_1 \in [A_1]_0^{\perp}, \ldots, c_n \in [A_n]_0^{\perp}$

 $(t|x_1=c_1,\ldots,x_n=c_n) \in [B]_0^{\perp}$ then:

• Assuming that $\alpha_i \in \llbracket \text{type of } \alpha_i \rrbracket^{\perp}$ (for i = 6, 9, 10, 11, 14, 15)

Adequacy (forcing mode)

lf

- \mathcal{E} : $X_1:A_1,\ldots,X_n:A_n\vdash t:B$ (in $\mathsf{PA}\omega^+$)
- $\bullet \ \rho \models \mathcal{E}^*, \quad c_1 \in \llbracket p_1 \Vdash A_1 \rrbracket_0^{\perp}, \ldots, c_n \in \llbracket p_n \Vdash A_n \rrbracket_0^{\perp}$

 $(t|x_1=c_1,\ldots,x_n=c_n)^* \in [(p_0p_1)\cdots p_n \Vdash B]_0^{\perp}$ then:

Conclusion



- This methodology applies to the forcing translation
 - A new abstract machine : the KFAM
 - Reminiscent from well known tricks of computer architecture (protection rings, virtual memory, hardware tracing, ...)
- How this computation model is used in particular cases of forcing?
- Use this methodology the other way around!
 - Deduce new logical translations from computation models borrowed to computer architecture, operating systems, ...