# Towards Intensionally More Expressive Systems for PTime

Stefan Schimanski

Department of Mathematics Ludwig-Maximilians-Universität München

Types 2006

- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(:) = 3$$
  
 $h(x;y) = c(:y,y,y)$   
 $f(n:) = rec(g,h,h)(n:)$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(:) = 3$$
  
 $h(x;y) = c(:y,y,y)$   
 $f(n:) = rec(g,h,h)(n:)$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;$ 



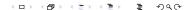
- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - ★ Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(:) = 3$$
  
 $h(x;y) = c(:y,y,y)$   
 $f(n:) = rec(g,h,h)(n:)$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;)$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;)$ 



- by relaxing linearity
- by combining different recursion schemes into one system
- by syntactical methods
  - finding decreasing measures for most general reduction sequences
  - considering sharing normalisation techniques
  - pointing out where special reduction strategies are essential
    - \* Folklore that BC needs special reduction strategy. What happens if we use proper sharing? Weiermann and Beckmann's example breaks down:

$$g(;) = 3$$
  
 $h(x;y) = c(;y,y,y)$   
 $f(n;) = rec(g,h,h)(n;)$ 



## First Example: linearity in LFPL

- LFPL = Hofmann's non-size-increasing term system
  - ► (affine) linearly typed
  - ▶ special ◊ type, seen as money
  - ▶ ♦ to be payed for list constructors

  - ightharpoonup  $\Rightarrow$  amount of  $\Diamond$  money doesn't increase during normalization
  - non-size-increasing iteration (list {step} base)
- Exactly LinSpace PTime algorithms representable
- Hofmann M.: Linear Types and Non-Size-Increasing Polynomial Time Computation. Logic in Computer Science (1998)

### Types

$$\sigma,\tau ::= \Diamond \mid B \mid \sigma \multimap \tau \mid \sigma \otimes \tau \mid \sigma \times \tau \mid L(\sigma)$$

#### Terms

$$s,t ::= x^{\tau} \mid c \mid \lambda x^{\tau}.t \mid \langle t,s \rangle \mid (t \ s) \mid \{t\}$$

#### Constructors

$$\begin{array}{ll} \mathbf{tt} \ \mathbf{ff} & B \\ \mathbf{cons}_{\tau} & \lozenge \multimap \tau \multimap \mathit{L}(\tau) \multimap \mathit{L}(\tau) \\ \mathbf{nil}_{\tau} & \mathit{L}(\tau) \\ \otimes_{\sigma,\tau} & \sigma \multimap \tau \multimap \sigma \otimes \tau \end{array}$$



## Complexity

## Theorem (Aehlig, Schwichtenberg)

For any typed term t there is a polynomial  $\vartheta(t)$  such that the length of a (special) reduction sequence is bounded by  $\vartheta(t)(|FV(t)|)$ .

#### Proof.

- ullet explicit definition of  $\vartheta(t)$  by recursion on t.
- $\vartheta(t)(N)$  decreases every conversion step. . . if  $\mathcal{L}(I) \leq N$  for every occurring list I
- linear typing  $\Rightarrow |FV(t)|$  doesn't increase
- and  $\mathcal{L}(I) \leq |FV(t)|$ .



## Restricted non-linearity in LFPL

Insertion Sort

```
\begin{aligned} &\operatorname{insert}(a,[]) = [a] \\ &\operatorname{insert}(a,b :: I) = \operatorname{if} \ a \leq b \ \operatorname{then} \ a :: b :: I \ \operatorname{else} \ b : \operatorname{insert}(a,I) \\ &\operatorname{sort}([]) = [] \\ &\operatorname{sort}(a :: I) = \operatorname{insert}(a,\operatorname{sort}(I)) \end{aligned}
```

Not linear:

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

• Intuition suggests:  $p^{\tau \to B} \in PTIME$  does not harm



## Restricted non-linearity in LFPL

Insertion Sort

```
\begin{aligned} &\operatorname{insert}(a,[]) = [a] \\ &\operatorname{insert}(a,b :: I) = \operatorname{if} \ a \leq b \ \operatorname{then} \ a :: b :: I \ \operatorname{else} \ b : \operatorname{insert}(a,I) \\ &\operatorname{sort}([]) = [] \\ &\operatorname{sort}(a :: I) = \operatorname{insert}(a,\operatorname{sort}(I)) \end{aligned}
```

Not linear:

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

ullet Intuition suggests:  $p^{ au o B}\in PTIME$  does not harm



- Operator  $\delta p$  of type  $\sigma \to (B \otimes \sigma)$  for  $p^{\sigma \to B}$ .
- Conversion  $(\delta p \ s) \mapsto (p \ s) \otimes s$
- Then

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

means

$$((\delta p \ x) \ \lambda y, z.(y \ \langle (f \ z), (g \ z) \rangle))$$

- This destroys linearity ⇒ only occurrences of variables count, not their names
- no easy measure for size anymore by counting variables, or even occurrences



- Operator  $\delta p$  of type  $\sigma \to (B \otimes \sigma)$  for  $p^{\sigma \to B}$ .
- Conversion  $(\delta p \ s) \mapsto (p \ s) \otimes s$
- Then

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

means

$$((\delta p \times) \lambda y, z.(y \langle (f z), (g z) \rangle))$$

- This destroys linearity ⇒ only occurrences of variables count, not their names
- no easy measure for size anymore by counting variables, or even occurrences



- Operator  $\delta p$  of type  $\sigma \to (B \otimes \sigma)$  for  $p^{\sigma \to B}$ .
- Conversion  $(\delta p \ s) \mapsto (p \ s) \otimes s$
- Then

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

means:

$$((\delta p \ x) \ \lambda y, z.(y \ \langle (f \ z), (g \ z) \rangle))$$

- This destroys linearity ⇒ only occurrences of variables count, not their names
- no easy measure for size anymore by counting variables, or even occurrences



- Operator  $\delta p$  of type  $\sigma \to (B \otimes \sigma)$  for  $p^{\sigma \to B}$ .
- Conversion  $(\delta p \ s) \mapsto (p \ s) \otimes s$
- Then

if 
$$p(x)$$
 then  $f(x)$  else  $g(x)$ 

means:

$$((\delta p \ x) \ \lambda y, z.(y \ \langle (f \ z), (g \ z) \rangle))$$

- This destroys linearity ⇒ only occurrences of variables count, not their names
- no easy measure for size anymore by counting variables, or even occurrences



#### **Terms**

$$s, t ::= x^{\tau} \mid c \mid \lambda x^{\tau}.t \mid \langle t, s \rangle \mid (t \mid s) \mid \{t\} \mid \delta f$$

#### Conversions

$$(\lambda x^{\tau}.t \ s) \mapsto t[s/x]$$

$$(s \otimes t \ r) \mapsto ((r \ s) \ t)$$

$$((cons \ d \ v \ x) \ \{h\} \ g) \mapsto (h \ d \ v \ (x \ \{h\} \ g))$$

$$(nil \ \{h\} \ g) \mapsto g$$

$$(tt \ \langle s, t \rangle) \mapsto s$$

$$(ff \ \langle s, t \rangle) \mapsto t$$

$$(\delta f \ t) \mapsto (f \ t) \otimes t$$



#### Quasi-linear typing rules

$$\frac{\Gamma, x^{\tau} \vdash x^{\tau}}{\Gamma, x^{\tau} \vdash x^{\tau}} (\text{Var}) \quad \frac{c \text{ of type } \tau}{\Gamma \vdash c^{\tau}} (\text{Const})$$

$$\frac{\Gamma, x^{\sigma} \vdash t^{\tau}}{\Gamma \vdash (\lambda x^{\sigma}.t)^{\sigma \to \tau}} (-\circ^{+}) \quad \frac{\Lambda_{1} \vdash t^{\sigma \to \tau}}{\Lambda_{1}, \Lambda_{2} \vdash (t s)^{\tau}} (-\circ^{-})$$

$$\frac{\Lambda_{1} \vdash t^{\rho \otimes \tau}}{\Lambda_{1}, \Lambda_{2} \vdash (t s)^{\sigma}} (\otimes^{-})$$

$$\frac{\Lambda_{1} \vdash t^{\rho \otimes \tau}}{\Lambda_{1}, \Lambda_{2} \vdash (t s)^{\sigma}} (\times^{-})$$

$$\frac{\Lambda \vdash t^{L(\tau)}}{\Lambda \vdash (t \{h\})^{\sigma \to \sigma}} (L(\tau)^{-})$$

$$\frac{\Gamma \vdash f^{\sigma \to \tau}}{\Gamma \vdash \delta f^{\sigma \to \tau \otimes \sigma}} (\delta^{+})$$

4D + 4D + 4E + 4E + 990

#### Quasi-linear typing rules

$$\frac{\Gamma, x^{\tau}; \Lambda \vdash x^{\tau}}{\Gamma; \Lambda, x^{\sigma} \vdash t^{\tau}} (\text{Var}) \quad \frac{c \text{ of type } \tau}{\Gamma; \Lambda \vdash c^{\tau}} (\text{Const})$$

$$\frac{\Gamma; \Lambda, x^{\sigma} \vdash t^{\tau}}{\Gamma; \Lambda \vdash (\lambda x^{\sigma}.t)^{\sigma \multimap \tau}} (\multimap^{+}) \quad \frac{\Gamma; \Lambda_{1} \vdash t^{\sigma \multimap \tau}}{\Gamma; \Lambda_{1}, \Lambda_{2} \vdash (t s)^{\tau}} (\multimap^{-})$$

$$\frac{\Gamma; \Lambda_{1} \vdash t^{\rho \otimes \tau}}{\Gamma; \Lambda_{1}, \Lambda_{2} \vdash (t s)^{\sigma}} (\otimes^{-})$$

$$\frac{\Gamma; \Lambda \vdash t^{L(\tau)}}{\Gamma; \Lambda \vdash (t \{h\})^{\sigma \multimap \sigma}} (L(\tau)^{-})$$

$$\frac{\Gamma; \Lambda \vdash f^{\sigma \multimap \tau}}{\Gamma; \Lambda \vdash \delta f^{\sigma \multimap \tau \otimes \sigma}} (\delta^{+})$$

#### Quasi-linear typing rules

$$\frac{c \text{ of type } \tau}{\Gamma, x^{\tau}; \Lambda \vdash x^{\tau}} (\text{Var}) \qquad \frac{c \text{ of type } \tau}{\Gamma; \Lambda \vdash c^{\tau}} (\text{Const}) \qquad \frac{\Gamma; \Lambda, x^{\tau} \vdash x^{\tau}}{\Gamma; \Lambda, x^{\tau} \vdash x^{\tau}} (\text{Var}_{I}) \qquad \frac{\Gamma; \Lambda, x^{\sigma} \vdash t^{\tau}}{\Gamma; \Lambda, x^{\sigma} \vdash t^{\tau}} (-\circ^{+}) \qquad \frac{\Gamma; \Lambda_{1} \vdash t^{\sigma \multimap \sigma}}{\Gamma; \Lambda_{1}, \Lambda_{2} \vdash (t s)^{\tau}} (-\circ^{-}) \qquad \frac{\Gamma; \Lambda_{1} \vdash t^{\rho \otimes \tau}}{\Gamma; \Lambda_{1}, \Lambda_{2} \vdash (t s)^{\sigma}} (\otimes^{-}) \qquad \frac{\Gamma; \Lambda_{1} \vdash t^{\rho \otimes \tau}}{\Gamma; \Lambda \vdash t^{L(\tau)}} \qquad \emptyset; \emptyset \vdash h^{\lozenge \multimap \tau \multimap \sigma \multimap \sigma}} (L(\tau)^{-}) \qquad \frac{\Gamma; \Lambda \vdash t^{L(\tau)}}{\Gamma; \Lambda \vdash \delta f^{\sigma \multimap \tau \otimes \sigma}} (\delta^{+})$$



## Reduction strategy matters

- Assume  $p = \lambda x^B$ .tt
- and  $t = \lambda I^{L(\tau)}.(I \{\lambda \Diamond, v, x^B.((\delta p \ x) \ \lambda y, z.y)\} \ \mathrm{tt}).$
- Then the following reduction sequence is possible:

$$(t (cons \lozenge_1 d_1 \dots (cons \lozenge_n d_n nil)))$$

$$\rightarrow \dots \rightarrow \underbrace{(\delta p (\delta p \dots tt).1).1}_{\text{depth } n}$$

• Applying all n  $\delta p$  conversions leads to  $2^{(n-1)}$  sub-terms of the form  $(\lambda x^B.\text{tt }t)$ , i.e. an **exponential complexity**.

## Reduction strategy matters

- Assume  $p = \lambda x^B$ .tt
- and  $t = \lambda I^{L(\tau)}.(I \{\lambda \Diamond, v, x^B.((\delta p \ x) \ \lambda y, z.y)\} \ \mathrm{tt}).$
- Then the following reduction sequence is possible:

$$(t (cons \lozenge_1 d_1 \dots (cons \lozenge_n d_n nil)))$$

$$\rightarrow \dots \rightarrow \underbrace{(\delta p (\delta p \dots tt).1).1}_{\text{depth } n}$$

• Applying all  $n \delta p$  conversions leads to  $2^{(n-1)}$  sub-terms of the form  $(\lambda x^B.\text{tt }t)$ , i.e. an **exponential complexity**.

10 / 17

## Reduction strategy matters

- Assume  $p = \lambda x^B$ .tt
- and  $t = \lambda I^{L(\tau)}.(I \{\lambda \Diamond, v, x^B.((\delta p \ x) \ \lambda y, z.y)\} \ \mathrm{tt}).$
- Then the following reduction sequence is possible:

$$(t (cons \lozenge_1 d_1 \dots (cons \lozenge_n d_n nil)))$$

$$\rightarrow \dots \rightarrow \underbrace{(\delta p (\delta p \dots tt).1).1}_{\text{depth } n}$$

• Applying all n  $\delta p$  conversions leads to  $2^{(n-1)}$  sub-terms of the form  $(\lambda x^B.\text{tt }t)$ , i.e. an **exponential complexity**.



## "Healthy" reduction strategy

- Reason for exponential growth: duplication of remaining "work"
- "one  $\delta p$  after the other"
- ullet  $\Rightarrow$  Only convert  $\delta p$  if the argument is in normal form
- or using sharing no special reduction strategy at all

## "Healthy" reduction strategy

- Reason for exponential growth: duplication of remaining "work"
- "one  $\delta p$  after the other"
- ullet  $\Rightarrow$  Only convert  $\delta p$  if the argument is in normal form
- or using sharing no special reduction strategy at all

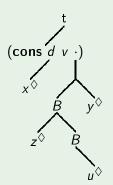
## "Healthy" reduction strategy

- Reason for exponential growth: duplication of remaining "work"
- ullet "one  $\delta p$  after the other"
- ullet  $\Rightarrow$  Only convert  $\delta p$  if the argument is in normal form
- or using sharing no special reduction strategy at all

- Inductive predicates to render interaction of variables
  - $ightharpoonup c_t(s)$  smallest "passive" super-term or t itself
  - $\triangleright$   $v \succ_t x$  between subterms  $v \trianglelefteq t$  and free variables  $x \in FV(t)$
  - ▶  $x \prec \succ_t y$  between  $x, y \in FV(t)$

## Interacting variables

## Example



- x and y interact.
- z and u do not interact.

## Interacting variables through $\lambda$

## Example

$$s = (\lambda x^{\lozenge}.w \ y^{\lozenge})$$

$$B$$

$$Cons \ d \ v \cdot)$$

$$z^{\lozenge} \ x^{\lozenge}$$

- Inductive predicates to render interaction of variables
  - $ightharpoonup c_t(s)$  smallest "passive" super-term or t itself
  - ▶  $v \succ_t x$  between subterms  $v \subseteq t$  and free variables  $x \in FV(t)$
  - ▶  $x \prec \succ_t y$  between  $x, y \in FV(t)$

  - $\exists c(c \succ_t x \land c \succ_t y) \rightarrow x \prec \succ_t y.$
- Induces equivalence relation over variables which reflects non-linearity
- every list resides completely in one class
- classes do not grow, only new classes are created
- new size measure: size of those equivalence classes
- ⇒ strong normalisation with sharing in PTime



- Inductive predicates to render interaction of variables
  - $ightharpoonup c_t(s)$  smallest "passive" super-term or t itself
  - ▶  $v \succ_t x$  between subterms  $v \subseteq t$  and free variables  $x \in FV(t)$
  - ▶  $x \prec \succ_t y$  between  $x, y \in FV(t)$
- Induces equivalence relation over variables which reflects non-linearity
- every list resides completely in one class
- classes do not grow, only new classes are created
- new size measure: size of those equivalence classes
- ⇒ strong normalisation with sharing in PTime



- Inductive predicates to render interaction of variables
  - $ightharpoonup c_t(s)$  smallest "passive" super-term or t itself
  - $\triangleright$   $v \succ_t x$  between subterms  $v \leq t$  and free variables  $x \in FV(t)$
  - ▶  $x \prec \succ_t y$  between  $x, y \in FV(t)$
  - $c_t(x) \succ_t x$
  - 2  $z \prec \succ_s x \land \lambda z^{\tau}.s \leq t \rightarrow c_t(\lambda z^{\tau}.s) \succ_t x$
  - $\exists c(c \succ_t x \land c \succ_t y) \rightarrow x \prec \succ_t y.$
- Induces equivalence relation over variables which reflects non-linearity
- every list resides completely in one class
- classes do not grow, only new classes are created
- new size measure: size of those equivalence classes
- ullet  $\Rightarrow$  strong normalisation with sharing in PTime



- LFPL originally developed to inject new base functions into BC
- What happens when adding nsi-It to BC directly?
- i.e. nsi-Iteration over incomplete/safe terms
- Naive:

⇒ in conflict with duplication of safe variables

- destroys sharing and separation of safe lists
- more precise analysis (e.g. using Gol) can characterise sane cases

- LFPL originally developed to inject new base functions into BC
- What happens when adding nsi-It to BC directly?
- i.e. nsi-Iteration over incomplete/safe terms
- Naive:

- ⇒ in conflict with duplication of safe variables
- destroys sharing and separation of safe lists
- more precise analysis (e.g. using Gol) can characterise sane cases

- LFPL originally developed to inject new base functions into BC
- What happens when adding nsi-It to BC directly?
- i.e. nsi-Iteration over incomplete/safe terms
- Naive:

- ⇒ in conflict with duplication of safe variables
- destroys sharing and separation of safe lists
- more precise analysis (e.g. using Gol) can characterise sane cases

- LFPL originally developed to inject new base functions into BC
- What happens when adding nsi-It to BC directly?
- i.e. nsi-Iteration over incomplete/safe terms
- Naive:

$$\begin{array}{lll} \mathsf{cat} &=& \lambda a, b. lt(a) (\lambda x, \diamondsuit, t, y. (\mathsf{cons} \ x \diamondsuit y)) b \\ \mathsf{exp}_1 &=& \lambda a. Rec((\mathsf{cons} \ 1 \diamondsuit \mathsf{nil})) (\lambda x, t, y. (\mathsf{cat} \ y \ y)) a \\ \mathsf{exp}_2 &=& \lambda a. Rec(\lambda z. (z, z) (\mathsf{cons} \ 1 \diamondsuit \mathsf{nil})) \\ &&&& (\lambda x, t, y. (\lambda z. (z, z) \ (\mathsf{cat} \ (\mathit{fst} \ y) \ (\mathit{snd} \ y)))) a \\ \end{array}$$

- ⇒ in conflict with duplication of safe variables
- destroys sharing and separation of safe lists
- more precise analysis (e.g. using Gol) can characterise sane cases



#### Conclusion

- Known systems can be extended
   ... but syntactical methods get much more complicated.
- Forcing reduction strategies is a tool to avoid "bad" behaviour
   ... though often sharing can remove those constraints.
- Hence: How essential are sharing or reduction strategies?
   LLL does not need that, more careful about duplication.
- Combination of recursion schemes not explored very much yet
- But fruitful to better understand dynamics of normalisation in PTime systems
  - ... e.g. importance of separation of safe data in BC.