Proving with Computer Assistance

Herman Geuvers

Lecture: Normalization for $\lambda\!\!\rightarrow$ and $\lambda 2$

Normalization of β

- Normal form A term M is in NF if there is no reduction step from M: $\neg\exists P(M \to_\beta P)$
- Weak Normalization A term M is WN if there is a reduction $M \to_\beta M_1 \to_\beta M_2 \to_\beta \ldots \to_\beta M_n$ with $M_n \in$ NF.
- Strong Normalization A term M is SN if there are no infinite reductions starting from M $\neg\exists(P_i)_{i\in\mathbb{N}}(M = P_0 \to_\beta P_1 \to_\beta P_2 \to_\beta \ldots)$.

We can give inductive definitions of NF, WN and SN.

Recap: relation between NF, WN and SN

Intermezzo: different definitions of "strong normalization"

M is SN if there are no infinite reductions starting from M

$\neg\exists(P_i)_{i\in\mathbb{N}}(M = P_0 \to_\beta P_1 \to_\beta P_2 \to_\beta \ldots)$

$\iff$ (classically) all β-reductions from M lead to a normal form

$\overset{??}{\iff}$ there is an upperbound k on the length of β-reductions from M.

Define $M \in \mathrm{SN}'$ as

$$\exists k \forall n, \forall P_1, \ldots, P_n (M = P_1 \to \ldots \to P_n) \implies n < k$$

- $M \in \mathrm{SN}' \implies M \in \mathrm{SN}$. (clearly, if there is a bound, then there is no infinite reduction)
- $M \in \mathrm{SN} \implies M \in \mathrm{SN}'$??
  Not in general for rewriting systems, but it holds for λ-calculus, because reduction in λ-calculus is finitely branching.

Proving (weak/strong) normalization of β

SN (or WN) for $\lambda\!\!\rightarrow$ cannot be proved by induction on the derivation

$$\frac{\Gamma \vdash M : \sigma \to \tau \qquad \Gamma \vdash N : \sigma}{\Gamma \vdash M\,N : \tau}$$

IH: M is SN and N is SN. So $M\,N$ is SN ??

No, e.g. $M = \lambda x.x\,x$, $N = \lambda x.x\,x$

Similarly for WN, the immediate induction proof fails.

We need an "induction loading": prove a stronger property that implies SN

Normalization of β for $\lambda\!\!\rightarrow$

Note:

- Terms may get larger under reduction
  $(\lambda f.\lambda x.f(fx))P \rightarrow_\beta \lambda x.P(Px)$
- Redexes may get multiplied under reduction.
  $(\lambda f.\lambda x.f(fx))((\lambda y.M)Q) \rightarrow_\beta \lambda x.((\lambda y.M)Q)(((\lambda y.M)Q)x)$
- New redexes may be created under reduction.
  $(\lambda f.\lambda x.f(fx))(\lambda y.N) \rightarrow_\beta \lambda x.(\lambda y.N)((\lambda y.N)x)$

First: Weak Normalization

- Weak Normalization: there is a reduction sequence that terminates,
- Strong Normalization: all reduction sequences terminate.

Weak Normalization

General property for (untyped) λ-calculus:
There are three ways in which a "new" β-redex can be created.

- Creation

$$(\lambda x. \ldots x\, P \ldots)(\lambda y.Q) \to_\beta \ldots (\lambda y.Q)P \ldots$$

- Multiplication

$$(\lambda x. \ldots x \ldots x \ldots)((\lambda y.Q)R) \to_\beta \ldots (\lambda y.Q)R \ldots (\lambda y.Q)R \ldots$$

- Identity

$$(\lambda x.x)(\lambda y.Q)R \to_\beta (\lambda y.Q)R$$

Weak Normalization

Proof originally from Turing, first published by Gandy (1980).

Definition

The height (or order) of a type $h(\sigma)$ is defined by

- $h(\alpha) := 0$
- $h(\sigma_1 \to \ldots \to \sigma_n \to \alpha) := \max(h(\sigma_1), \ldots, h(\sigma_n)) + 1.$

NB [Exercise] This is the same as defining

- $h(\sigma \to \tau) := \max(h(\sigma) + 1, h(\tau))$.

Definition

The height of a redex $(\lambda x{:}\sigma.P)Q$ is the height of the type of $\lambda x{:}\sigma.P$

Weak Normalization

Definition

We give a measure m to the terms by defining $m(N) := (h(N), \#N)$ with

- $h(N) =$ the maximum height of a redex in N,
- $\#N =$ the number of redexes of height $h(N)$ in N.

The measures of terms are ordered lexicographically:

$$(h_1, m) <_l (h_2, n) \text{ iff } h_1 < h_2 \text{ or } (h_1 = h_2 \text{ and } m < n).$$

Theorem: Weak Normalization

If P is a typable term in $\lambda\!\!\rightarrow$, then there is a terminating reduction starting from P.

Proof

Pick a redex of height $h(P)$ that does not contain any other redex of height $h(P)$. [Note that this is always possible!]
Contract this redex, to obtain Q.

Claim: This does not create a new redex of height $h(P)$.
This is the important step. [Exercise: check this; use the three ways in which new redexes can be created.]

So $m(P) >_l m(Q)$, because $m(P) = (h(P), \#P)$ and either

- the number of redexes of height $h(P)$ has decreased by 1, and then $m(Q) = (h(P), \#P - 1)$, or
- there are no redexes of height $h(P)$ left, and then $m(Q) = (h(Q), n)$, with $h(P) > h(Q)$ for some n.

As there are no infinitely decreasing $<_l$ sequences, this process must terminate and then we have arrived at a normal form.

Strong Normalization for $\lambda\!\!\rightarrow$ à la Curry

This is proved by constructing a model of $\lambda\!\!\rightarrow$.
Method originally due to Tait (1967); also direct "arithmetical" methods exist, that use a decreasing ordering (David 2001, David & Nour)

Definition

- $[\![\alpha]\!] := \mathsf{SN}$ (the set of strongly normalizing λ-terms).
- $[\![\sigma\!\rightarrow\!\tau]\!] := \{M \mid \forall N \in [\![\sigma]\!](MN \in [\![\tau]\!])\}$.

Lemma

1. $xN_1 \ldots N_k \in [\![\sigma]\!]$ for all x, σ and $N_1, \ldots, N_k \in \mathsf{SN}$.
2. $[\![\sigma]\!] \subseteq \mathsf{SN}$
3. If $M[x := N]\vec{P} \in [\![\sigma]\!]$, $N \in \mathsf{SN}$, then $(\lambda x.M)N\vec{P} \in [\![\sigma]\!]$.

Lemma for Strong Normalization

Lemma cases (1) and (2)

1. $xN_1 \ldots N_k \in [\![\sigma]\!]$ for all x, σ and $N_1, \ldots, N_k \in \mathsf{SN}$.
2. $[\![\sigma]\!] \subseteq \mathsf{SN}$

Proof: Simultaneously by induction on σ.

Lemma for Strong Normalization

Lemma case (3)

3. If $M[x := N]\vec{P} \in [\![\sigma]\!]$, $N \in \mathsf{SN}$, then $(\lambda x.M)N\vec{P} \in [\![\sigma]\!]$.

Proof: By induction on σ.

Proposition

$$\left.\begin{array}{r} x_1{:}\tau_1, \ldots, x_n{:}\tau_n \vdash M : \sigma \\ N_1 \in [\![\tau_1]\!], \ldots, N_n \in [\![\tau_n]\!] \end{array}\right\} \Rightarrow M[x_1 := N_1, \ldots x_n := N_n] \in [\![\sigma]\!]$$

Proof By induction on the derivation of $\Gamma \vdash M : \sigma$. (Using (3) of the previous Lemma.)

Proposition

$$\left.\begin{array}{r} x_1{:}\tau_1, \ldots, x_n{:}\tau_n \vdash M : \sigma \\ N_1 \in [\![\tau_1]\!], \ldots, N_n \in [\![\tau_n]\!] \end{array}\right\} \Rightarrow M[x_1 := N_1, \ldots x_n := N_n] \in [\![\sigma]\!]$$

Corollary $\lambda\!\rightarrow$ is SN

Proof By taking $N_i := x_i$ in the Proposition. (That can be done, because $x_i \in [\![\tau_i]\!]$ by (1) of the Lemma.)
Then $M \in [\![\sigma]\!] \subseteq$ SN, using (2) of the Lemma. QED

Exercise Verify the details of the Strong Normalization proof. (That is, prove the missing details of the Lemma and the Proposition.)

Consistency

Normalization (weak or strong) imples logical consistency of the type theory: there is a type A that has no closed inhabitant:

$$\neg\exists M(\vdash M : A)$$

Proof.

A little bit on semantics

$\lambda\!\!\rightarrow$ has a simple set-theoretic model. Given sets $[\![\alpha]\!]$ for type variables α, define

$$[\![\sigma\!\rightarrow\!\tau]\!] := [\![\tau]\!]^{[\![\sigma]\!]} \text{ (set theoretic function space } [\![\sigma]\!] \rightarrow [\![\tau]\!])$$

If any of the base sets $[\![\alpha]\!]$ is infinite, then there are higher and higher (uncountable) cardinalities among the $[\![\sigma]\!]$
There are smaller models, e.g.

$$[\![\sigma\!\rightarrow\!\tau]\!] := \{f \in [\![\sigma]\!] \rightarrow [\![\tau]\!] | f \text{ is definable}\}$$

where definability means that it can be constructed in some formal system. This restricts the collection to a countable set.
For example

$$[\![\sigma\!\rightarrow\!\tau]\!] := \{f \in [\![\sigma]\!] \rightarrow [\![\tau]\!] | f \text{ is } \lambda\text{-definable}\}$$

$\lambda 2$

Church style:

$$\frac{\Gamma \vdash M : \sigma}{\Gamma \vdash \lambda\alpha.M : \forall\alpha.\sigma}\ \alpha \notin \mathsf{FV}(\Gamma) \qquad \frac{\Gamma \vdash M : \forall\alpha.\sigma}{\Gamma \vdash M\tau : \sigma[\alpha := \tau]}\ \text{for } \tau \text{ a } \lambda 2\text{-type}$$

Curry style:

$$\frac{\Gamma \vdash M : \sigma}{\Gamma \vdash M : \forall\alpha.\sigma}\ \alpha \notin \mathsf{FV}(\Gamma) \qquad \frac{\Gamma \vdash M : \forall\alpha.\sigma}{\Gamma \vdash M : \sigma[\alpha := \tau]}\ \text{for } \tau \text{ a } \lambda 2\text{-type}$$

Strong Normalization of β for $\lambda 2$

- For $\lambda 2$ a la Church, there are two kinds of β-reductions:
  - $(\lambda x{:}\sigma.M)P \to_\beta M[x := P]$ term reduction
  - $(\lambda \alpha.M)\tau \to_\beta M[\alpha := \tau]$ type reduction
- The second doesn't do any harm, so we can just look at $\lambda 2$ à la Curry
  More precisely:
  - type reduction is terminating
  - if there is an infinite combined term reduction / type reduction path in $\lambda 2$ a la Church, then there is an infinite term reduction path in $\lambda 2$ a la Curry.

Strong Normalization of β for $\lambda 2$ a la Curry

Recall the proof for $\lambda\!\!\rightarrow$:

- $[\![\alpha]\!] := \mathsf{SN}$.
- $[\![\sigma\!\rightarrow\!\tau]\!] := \{M \mid \forall N \in [\![\sigma]\!](MN \in [\![\tau]\!])\}$.

Question:
How to define $[\![\forall\alpha.\sigma]\!]$??

$$[\![\forall\alpha.\sigma]\!] := \Pi_{X \in U}[\![\sigma]\!]_{\alpha:=X}??$$

Interpretation of types

Question: How to define $[\![\forall\alpha.\sigma]\!]$??

$$[\![\forall\alpha.\sigma]\!] := \Pi_{X\in U}[\![\sigma]\!]_{\alpha:=X}??$$

- What should U be?
  The collection of "all possible interpretations" of types (?)
- $\Pi_{X\in U}[\![\sigma]\!]_{\alpha:=X}$ gets too big: $\mathrm{card}(\Pi_{X\in U}[\![\sigma]\!]_{\alpha:=X}) > \mathrm{card}(U)$

Girard:

- $[\![\forall\alpha.\sigma]\!]$ should be small

$$\bigcap_{X\in U}[\![\sigma]\!]_{\alpha:=X}$$

- Characterization of U.

Saturated sets

$U := \mathrm{SAT}$, the collection of saturated sets of (untyped) λ-terms.
$X \subset \Lambda$ is saturated if

- $xP_1 \dots P_n \in X$ (for all $x \in \mathrm{Var}$, $P_1, \dots, P_n \in \mathrm{SN}$)
- $X \subseteq \mathrm{SN}$
- If $M[x := N]\vec{P} \in X$ and $N \in \mathrm{SN}$, then $(\lambda x.M)N\vec{P} \in X$.

Let $\rho : \mathrm{TVar} \to \mathrm{SAT}$ be a valuation of type variables.
Define the interpretation of types $[\![\sigma]\!]_\rho$ as follows.

- $[\![\alpha]\!]_\rho := \rho(\alpha)$
- $[\![\sigma \to \tau]\!]_\rho := \{M | \forall N \in [\![\sigma]\!]_\rho (MN \in [\![\tau]\!]_\rho)\}$
- $[\![\forall \alpha.\sigma]\!]_\rho := \cap_{X \in \mathrm{SAT}} [\![\sigma]\!]_{\rho, \alpha := X}$

Soundness property

Proposition

$$x_1 : \tau_1, \ldots, x_n : \tau_n \vdash M : \sigma \Rightarrow M[x_1 := P_1, \ldots, x_n := P_n] \in [\![\sigma]\!]_\rho$$

for all valuations ρ and $P_1 \in [\![\tau_1]\!]_\rho, \ldots, P_n \in [\![\tau_n]\!]_\rho$

Proof

By induction on the derivation of $\Gamma \vdash M : \sigma$.

Corollary $\lambda 2$ is SN

(Proof: take P_1 to be x_1, ..., P_n to be x_n.)

A little bit on semantics

$\lambda 2$ does not have a set-theoretic model! [Reynolds]

Theorem: If

$$[\![\sigma \to \tau]\!] := [\![\tau]\!]^{[\![\sigma]\!]} \text{ (set theoretic function space)}$$

then $[\![\sigma]\!]$ is a singleton set for every σ.

So: in a $\lambda 2$-model, $[\![\sigma \to \tau]\!]$ must be 'small'.