Substitution Theorem

Makoto Tatsuta (National Institute of Informatics)

(joint work with Mariangiola Dezani-Ciancaglini (Torino University))

Seminar
School of Computer Science and Engineering
Seoul National University
March 16, 2007

Introduction

Results:

Theorem (Substitution Theorem) If $M[x_1 := X, \ldots, x_n := X] \in$ SN for all $X \in$ SN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ SN for all $X_1, \ldots, X_n \in$ SN.

$MXXX \in$ SN for all $X \in$ SN implies $MXYZ \in$ SN for all $X, Y, Z \in$ SN

M persistently strongly normalizing ($M \in$ PSN) iff $MN_1 \ldots N_n \in$ SN for all n and $N_1, \ldots, N_n \in$ SN

Theorem. $M \in$ PSN iff $\Gamma_\omega \vdash M : \omega$ in the type theory $\mathcal{HL}$ where $\Gamma_\omega = \{x : \omega \mid x \in \mathit{Var}\}$.

This solves the open question in Dezani, Honsell and Motohama (TCS2005)

Ideas:
- adjacent controls
- Klop's calculus
- inductive definitions of PSN terms

Related Work

Dezani, Honsell and Motohama (TCS2005)

$M \in PWN$ iff $MN_1 \dots N_n \in WN$ for all n and $N_1, \dots, N_n \in WN$
$M \in PHN$ iff $MN_1 \dots N_n \in HN$ for all n and $N_1, \dots, N_n$
$M \in PWHN$ iff $MN_1 \dots N_n \in WHN$ for all n and $N_1, \dots, N_n$

Characterization by a type theory with intersection types

Eg. $M \in PWN$ iff $M : \omega$ in the type theory CDZ

Characterization of PSN is an open question

Weak Normalization

Theorem (Substitution Theorem for WN**).** If $M[x_i := X, x_j := X] \in$ WN for all i, j $(1 \leq i, j \leq n)$ for all $X \in$ WN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ WN for all $X_1, \ldots, X_n \in$ WN.

Eg. $M = x_1(yx_2)$.
$M[x_1 := X, x_2 = X] \in$ WN holds for all $X \in$ WN.
Moreover $M[x_1 := X_1, x_2 = X_2] \in$ WN holds for all $X_1, X_2 \in$ WN.

Proof Ideas:

Adjacent controls
- M has no adjacent controls for x, y
$\iff \quad M[x := X, y := Y] \in$ WN for all $X, Y \in$ WN

Controls

Eg. $M = x(\lambda z.zy)$ and $X = \lambda u.u\Delta$

$M[x := X] = (\lambda u.u\Delta)(\lambda z.zy) \to^* \Delta y$

- By putting an appropriate term into x and reducing it, we can put an arbitrary term into z.
- We will say that x controls z and write $x \rightsquigarrow z$.

More generally, $M = x(\lambda x_1.\ldots x_1(\lambda x_2.\ldots x_2(\lambda x_3.\ldots x_3 z)))$

$x \rightsquigarrow x_1$ and $x_1 \rightsquigarrow x_2$ and $x_2 \rightsquigarrow x_3$

$x \rightsquigarrow x_3$

Definition of Controls

x sometimes denotes an occurrence of the variable x in a term, which may be bound.

Let M be in β-normal form and the occurrences x and y be in M.

The relation $x \leadsto y$ in M is the reflexive transitive closure of the following relation $\leadsto_1$.
- if $x\overrightarrow{N}(\lambda\overrightarrow{t}y.L)$ is a subterm of M then $x \leadsto_1 y$ in M.

We say x controls y in M if $x \leadsto y$ in M.

We also call the pair (x, y) the control when $x \leadsto y$. We sometimes say the control $x \leadsto y$ to denote this (x, y).

If $x \leadsto y$, by putting an appropriate term into x in M and reducing it we can have an arbitrary term in y.

Example.
$x \leadsto z$ in $\lambda y.x(\lambda v.vvy)(\lambda t.t(\lambda uz.xz))$

Adjacent Controls

Adjacent variable occurrences
- two variable occurrences x and y in M are adjacent if M contains a subterm of the shape $x\vec{N}(\lambda\vec{u}.y\vec{L})$

Eg. xzy $\quad$ $x(\lambda z.y)$

x and y form an (indirect) application.

Adjacent controls
- two controls $x \rightsquigarrow z$ and $y \rightsquigarrow t$ in M are adjacent controls for x, y in M if z and t are adjacent in M.

Example.
v and y are adjacent in $\lambda w.x(\lambda v.vvy)(\lambda t.t(\lambda uz.xz))$.
$x \rightsquigarrow v$ and $y \rightsquigarrow y$ are adjacent controls in this term.

Since z and t form an (indirect) application, and x controls z and y controls t, we have some $X, Y \in$ WN so that $M[x := X, y := Y] \notin$ WN.

Let V be a set of free variables.
We say $x \rightsquigarrow u$ and $y \rightsquigarrow v$ are adjacent controls for V in M if $x, y \in V$.

Key Lemma for WN

Adjacent controls are essentially the same as adjacent replacement paths in [Dezani-Honsell-Motohama2005].

Lemma[Dezani-Honsell-Motohama2005]. Let $M \in \beta$-NF. If there are adjacent controls for x in M, then there is $X \in$ WN such that $M[x := X] \notin$ WN.

Lemma 1. If M is in β-NF and for all $X \in$ WN we have $M[x := X, y := X] \in$ WN then there are no adjacent controls for x, y in M.

Proof. Let M be $M[y := x]$ in the previous lemma.

Lemma 2 (Key Lemma). If $M \in \beta$-NF and there are no adjacent controls for $x_1, \ldots, x_n$ in M, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ WN for all $X_1, \ldots, X_n \in$ WN.

Substitution Theorem for WN

Theorem. If $M[x_i := X, x_j := X] \in$ WN for all i, j $(1 \leq i, j \leq n)$ for all $X \in$ WN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ WN for all $X_1, \ldots, X_n \in$ WN.

Proof.

Let $\vec{x}$ be $x_1, \ldots, x_n$ and N be the β-normal form of M. By Lemma 1 there are no adjacent controls for x_i, x_j $(1 \leq i, j \leq n)$ in N. Hence there are no adjacent controls for $\vec{x}$ in N. By Lemma 2, we have $N[\vec{x} := \vec{X}] \in$ WN for all $\vec{X} \in$ WN. Hence we have $M[\vec{x} := \vec{X}] \in$ WN for all $\vec{X} \in$ WN. $\square$

Strong Normalization

Theorem (Substitution Theorem) If $M[x_i := X, x_j := X] \in$ SN for all i, j $(1 \leq i, j \leq n)$ for all $X \in$ SN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ SN for all $X_1, \ldots, X_n \in$ SN.

Theorem (Substitution Theorem) If $M[x_1 := X, \ldots, x_n := X] \in$ SN for all $X \in$ SN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ SN for all $X_1, \ldots, X_n \in$ SN.

Difficulty:

- controls can be defined only for β-normal forms
- β-reduction preserves only WN

Idea:

- Klop's calculus

Klop's calculus 1

λ^*-calculus (Boudol, TLCA2003)

$S ::= x|\lambda x.S|SS|[S, S]$

S keeps an erased argument.

$[S, T_1, \ldots, T_n]$ and $[S, \overrightarrow{T}]$ are short for $[\ldots[[S, T_1], T_2], \ldots T_n]$

κ-reduction:

$[\lambda x.S, U_1, \ldots, U_n]T \to_\kappa [S[x := T], U_1, \ldots, U_n]$ if $x \in FV(S)$

$[\lambda x.S, U_1, \ldots, U_n]T \to_\kappa [S, U_1, \ldots, U_n, T]$ if $x \notin FV(S)$

Eg.

$(\lambda x.\underline{(\lambda yz.z)(xx)}z)\Delta \to_\kappa (\lambda x.[\lambda z.z, xx]z)\Delta \to_\kappa (\lambda x.[z, xx])\Delta \to_\kappa [z, \Delta\Delta]$.

Klop's calculus 2

SN^* the set of strongly normalizing terms for κ-reduction
WN^* the set of weakly normalizing terms for κ-reduction
SN the set of strongly normalizing λ-terms for β-reduction
Λ the set of λ-terms

Theorem[Boudol].
(1) $\mathsf{SN}^* = \mathsf{WN}^*$.
(2) $\mathsf{SN} \supseteq \mathsf{SN}^* \cap \Lambda$.

Theorem. $\mathsf{SN} \subseteq \mathsf{SN}^*$.

Combining these, we have $\mathsf{SN} = \mathsf{SN}^* \cap \Lambda$. That is, λ^* is conservative over λ with respect to strong normalization.

Klop's calculus enables us to use controls for SN case

Projections

The set $\mathcal{P}_a(S)$ of λ-terms simulates the behavior of S.

Fix a fresh variable a.
The set $\mathcal{P}_a(S)$ of projections of S:

$$\begin{array}{lcl} \mathcal{P}_a(x) & = & \{x\}, \\ \mathcal{P}_a(\lambda x.S) & = & \lambda x.\mathcal{P}_a(S), \\ \mathcal{P}_a(ST) & = & \mathcal{P}_a(S)\mathcal{P}_a(T), \\ \mathcal{P}_a([S,T]) & = & \mathcal{P}_a(S) \cup a\mathcal{P}_a(T). \end{array}$$

where $\mathcal{A}, \mathcal{B}$ sets of λ-terms

$$\begin{array}{l} \lambda x.\mathcal{A} = \{\lambda x.M | M \in \mathcal{A}\} \\ \mathcal{A}\mathcal{B} = \{MN | M \in \mathcal{A}, N \in \mathcal{B}\} \\ a\mathcal{A} = \{aM | M \in \mathcal{A}\}. \end{array}$$

Lemma.
(1) $M \in \mathcal{P}_a(S)$ and $N \in \mathcal{P}_a(T)$ imply $M[x := N] \in \mathcal{P}_a(S[x := T])$.
(2) If $S \in \mathsf{SN}^*$, then $\mathcal{P}_a(S) \subseteq \mathsf{SN}$.

Adjacent Controls for λ^*

Controls are defined by using projections in a similar way to Λ

Definition.
- We have a control $x \rightsquigarrow y$ in S
if we have the control $x \rightsquigarrow y$ in M for some $M \in \mathcal{P}_a(S)$.
- We have adjacent controls $x \rightsquigarrow u$ and $y \rightsquigarrow v$ in S
if we have the adjacent controls $x \rightsquigarrow u$ and $y \rightsquigarrow v$ in M for some $M \in \mathcal{P}_a(S)$.

Key Lemma and Theorem

Lemma 3. If $S \in \kappa$-NF and $S[x := X, y = X] \in \mathsf{SN}^*$ for all $X \in \beta$-NF then there are no adjacent controls for x, y in S.

Lemma 4 (Key Lemma). If there are no adjacent controls for $\overrightarrow{x}$ in a κ-normal form S, then $S[\overrightarrow{x} := \overrightarrow{X}] \in \mathsf{SN}^*$ for all $\overrightarrow{X} \in \mathsf{SN}^*$.

Theorem (Substitution Theorem) If $M[x_i := X, x_j := X] \in \mathsf{SN}$ for all i, j $(1 \leq i, j \leq n)$ for all $X \in \mathsf{SN}$, then $M[x_1 := X_1, \ldots, x_n := X_n] \in \mathsf{SN}$ for all $X_1, \ldots, X_n \in \mathsf{SN}$.

Proof. Let $M \to_\kappa^* K \in \kappa$-NF. By SN $\subseteq$ SN*, for all $X \in \beta$-NF we have $K[x_i := X, x_j := X] \in \mathsf{SN}^*$. By Lemma 3 there are no adjacent controls for x_i, x_j $(1 \leq i, j \leq n)$ in K. Hence there are no adjacent controls for $\vec{x}$. By Lemma 4, we have $K[\vec{x} := \vec{X}] \in \mathsf{SN}^*$ for all $\vec{X} \in \mathsf{SN}^*$. By Theorem (Boudol), $M[\vec{x} := \vec{X}] \in \mathsf{SN}$ for all $\vec{X} \in \mathsf{SN}$. $\square$

Inductive Definition of PSN

$$\frac{\lambda \overrightarrow{x}.N \in \mathsf{SN}^\sharp_n \ (\forall N \in \overrightarrow{N}) \quad lh(\overrightarrow{x}) = n \quad y \notin \overrightarrow{x}}{\lambda \overrightarrow{x}.y\overrightarrow{N} \in \mathsf{PSN}^\sharp}$$

$$\frac{\lambda \overrightarrow{x}.M[y := N]\overrightarrow{L} \in \mathsf{PSN}^\sharp \quad \lambda \overrightarrow{x}.N \in \mathsf{SN}^\sharp_n \quad lh(\overrightarrow{x}) = n}{\lambda \overrightarrow{x}.(\lambda y.M)N\overrightarrow{L} \in \mathsf{PSN}^\sharp}$$

$$\frac{\lambda \overrightarrow{x}.N \in \mathsf{PSN}^\sharp \ (\forall N \in \overrightarrow{N}) \quad lh(\overrightarrow{x}) = n \quad x \in \overrightarrow{x}}{\lambda \overrightarrow{x}.x\overrightarrow{N} \in \mathsf{SN}^\sharp_n}$$

$$\frac{\lambda \overrightarrow{x}.N \in \mathsf{SN}^\sharp_m \ (\forall N \in \overrightarrow{N}) \quad lh(\overrightarrow{x}) = m \quad y \notin \overrightarrow{x}}{\lambda \overrightarrow{x}.y\overrightarrow{N} \in \mathsf{SN}^\sharp_n}$$

$$\frac{\lambda \overrightarrow{x}.N \in \mathsf{SN}^\sharp_n \quad lh(\overrightarrow{x}) = n \quad lh(\overrightarrow{y}) > 0}{\lambda \overrightarrow{x}\,\overrightarrow{y}.N \in \mathsf{SN}^\sharp_n}$$

$$\frac{\lambda \overrightarrow{x}.M[y := N]\overrightarrow{L} \in \mathsf{SN}^\sharp_n \quad \lambda \overrightarrow{x}.N \in \mathsf{SN}^\sharp_m \quad lh(\overrightarrow{x}) = m}{\lambda \overrightarrow{x}.(\lambda y.M)N\overrightarrow{L} \in \mathsf{SN}^\sharp_n}$$

$M \in \mathsf{SN}_n$ iff $MX_1 \ldots X_n \in \mathsf{SN}$ for all $X_1, \ldots, X_n \in \mathsf{SN}$

Soundness and Completeness

Theorem. $\mathsf{PSN}^\sharp = \mathsf{PSN}$ and $\mathsf{SN}_n^\sharp = \mathsf{SN}_n$.

Lemma 5. If $\lambda\vec{x}.x\vec{N} \in \mathsf{SN}_n$ where $x \in \vec{x}$ and $lh(\vec{x}) = n$, then $\lambda\vec{x}.N_i \in \mathsf{PSN}$ for all $N_i \in \vec{N}$.

Proof. For arbitrary $\vec{X} \in \mathsf{SN}$ of length n, we have $(x\vec{N})[\vec{x} := \vec{X}] \in \mathsf{SN}$. Suppose $y \notin x\vec{N}$. By Substitution Theorem, $(y\vec{N})[\vec{x} := \vec{X}, y := Y] \in \mathsf{SN}$ for all $\vec{X}, Y \in \mathsf{SN}$. For $N_i \in \vec{N}$, we will show $(\lambda\vec{x}.N_i)\vec{X}\vec{Z} \in \mathsf{SN}$ for given $\vec{X}, \vec{Z} \in \mathsf{SN}$. Let Y be $\lambda\vec{m}.m_i\vec{Z}$. Then
$(y\vec{N})[\vec{x} := \vec{X}, y := Y] = (\lambda\vec{m}.m_i\vec{Z})\vec{N}[\vec{x} := \vec{X}] \to_\beta^* N_i[\vec{x} := \vec{X}]\vec{Z}$. Hence $N_i[\vec{x} := \vec{X}]\vec{Z} \in \mathsf{SN}$. Then $(\lambda\vec{x}.N_i)\vec{X}\vec{Z} \in \mathsf{SN}$. Therefore $\lambda\vec{x}.N_i \in \mathsf{PSN}$. $\Box$

Type Theory $\mathcal{HL}$

Types $\sigma ::= \omega | \varphi | \sigma \to \sigma | \sigma \cap \sigma$
- No type variables. Only two type constants ω, φ.

Type preorder

$$\sigma \leq \sigma \cap \sigma \qquad \sigma \cap \tau \leq \sigma \qquad \sigma \cap \tau \leq \tau$$

$$\sigma \leq \sigma',\ \tau \leq \tau' \Rightarrow \sigma \cap \sigma' \leq \tau \cap \tau'$$

$$\sigma' \leq \sigma,\ \tau \leq \tau' \Rightarrow \sigma \to \tau \leq \sigma' \to \tau'$$

$$(\sigma \to \tau) \cap (\sigma \to \zeta) \leq \sigma \to \tau \cap \zeta$$

$$\varphi \sim \omega \to \varphi \qquad \omega \sim \varphi \to \omega \qquad \omega \leq \varphi$$

$$\sigma \leq \sigma \qquad \sigma \leq \tau,\ \tau \leq \zeta\ \Rightarrow \sigma \leq \zeta$$

Typing rules

$$(\mathsf{Ax})\ \frac{(x{:}\sigma) \in \Gamma}{\Gamma \vdash x{:}\sigma} \qquad (\to \mathrm{I})\ \frac{\Gamma, x{:}\sigma \vdash M : \tau}{\Gamma \vdash \lambda x.M : \sigma \to \tau}$$

$$(\to \mathsf{E})\ \frac{\Gamma \vdash M : \sigma \to \tau \quad \Gamma \vdash N : \sigma}{\Gamma \vdash MN : \tau}$$

$$(\leq)\ \frac{\Gamma \vdash M : \sigma \quad \sigma \leq \tau}{\Gamma \vdash M : \tau} \qquad (\cap \mathrm{I})\ \frac{\Gamma \vdash M : \sigma \quad \Gamma \vdash M : \tau}{\Gamma \vdash M : \sigma \cap \tau}$$

Characterization Theorem of PSN

Theorem. Let $\Gamma_\omega = \{x : \omega | x \in Var\}$.
(1) $M \in \mathsf{SN}$ iff $\Gamma_\omega \vdash M : \varphi$.
(2) $M \in \mathsf{PSN}$ iff $\Gamma_\omega \vdash M : \omega$.

Proof of the claims (1) and (2) from the left hand side to the right hand side.

Assume $M \in \mathsf{SN}$ or $M \in \mathsf{PSN}$. We will show $\Gamma_\omega \vdash M : \varphi$ or $\Gamma_\omega \vdash M : \omega$. By inductive definitions of PSN and SN_n, we have $M \in \mathsf{SN}_0^\sharp$ or $M \in \mathsf{PSN}^\sharp$. By induction on the proof of $M \in \mathsf{SN}_0^\sharp$ or $M \in \mathsf{PSN}^\sharp$, we can show $\Gamma_\omega \vdash M : \varphi$ or $\Gamma_\omega \vdash M : \omega$. $\square$

Conclusion

Results:

Theorem (Substitution Theorem) If $M[x_1 := X, \ldots, x_n := X] \in$ SN for all $X \in$ SN, then $M[x_1 := X_1, \ldots, x_n := X_n] \in$ SN for all $X_1, \ldots, X_n \in$ SN.

Theorem. $M \in$ PSN iff $\Gamma_\omega \vdash M : \omega$ in the type theory $\mathcal{HL}$ where $\Gamma_\omega = \{x{:}\omega | x \in \mathit{Var}\}$.

This solves the open question in Dezani, Honsell and Motohama (TCS2005)

Ideas:

- adjacent controls
- Klop's calculus
- inductive definitions of PSN terms