Substitutions en λ-calcul simplement typé

Chantal KELLER

École Normale Supérieure de Lyon

11 septembre 2008

Plan 1/2

Plan 2/2

3 La catégorie des substitutions parallèles
- Substitutions parallèles
- Preuves

4 Substitutions héréditaires et décidabilité de la $\beta\eta$-équivalence
- $\beta\eta$-équivalence et formes normales
- Substitutions héréditaires
- Normalisation
- Preuves

Première partie I

Cadre et but du stage

Cadre

Stage de dix semaines dans le laboratoire de Programmation Fonctionnelle de l'École d'Informatique de Nottingham, en Angleterre.

But

λ-calcul :

- modèle de calcul introduit dans les années 1930
- dans lequel les substitutions jouent un rôle important

Exemple : β-réduction :

$$(\lambda x.f)\ t \to f[x := t]$$

But

Différentes approches pour les substitutions :

- ponctuelles $\neq$ parallèles ;
- "classiques" $\neq$ héréditaires ;
- $\ldots$ etc.

Ici, λ-calcul simplement typé avec :

- substitutions parallèles (en tant qu'objets de première classe) ;
- substitutions héréditaires implantées avec des substitutions ponctuelles.

Substitutions parallèles

- Point de vue élégant sur les substitutions ...
- ... ce qui permet de prouver facilement des propriétés dessus.

Ici : “Les substitutions parallèles forment une catégorie munies de produits finis.”

Substitutions héréditaires

- Type de substitutions qui est **structurellement** récursif ;
- $\Rightarrow$ on peut l'implanter de manière élégante en utilisant un assistant de preuves ...
- ... puis prouver des propriétés dessus.

Ici : “La $\beta\eta$-équivalence est décidable (pour les substitutions héréditaires).”

Assistant de preuves

Ces preuves ne sont pas (complètement) nouvelles, mais l'intérêt est de les implanter en utilisant un assistant de preuves : **Agda**.

- participe à l'intérêt de développer de tels outils ;
- donne des directions de développement ;
- donne un exemple d'utilisation ;
- fournit des bibliothèques de preuves pour faire de nouvelles preuves.

Deuxième partie II

Le λ-calcul simplement typé

Particularités

Modèle de calcul connu, avec quelques particularités dans notre cas :

- utilisation d'une syntaxe directement typée (pas de terme inutile) ;
- utilisation de la notation à la de Bruijn (pas de problème avec les noms des variables).

Types

Un seul type de base :

$$\frac{}{Ty : \mathrm{Set}} \qquad \frac{}{\circ : Ty}\ \mathrm{base} \qquad \frac{\sigma, \tau : Ty}{\sigma \rightarrow \tau : Ty}\ \mathrm{fun}$$

Contextes

Listes de types :

$$\frac{}{Con : \mathrm{Set}} \qquad \frac{}{\varepsilon : Con}\ \varepsilon \qquad \frac{\Gamma : Con \qquad \sigma : Ty}{\Gamma, \sigma : Con}\ \mathrm{ext}$$

Variables

"Numéros" dans un contexte donné :

$$\frac{\Gamma : \mathit{Con} \qquad \sigma : \mathit{Ty}}{\mathit{Var}\ \Gamma\ \sigma : \mathrm{Set}}$$

$$\frac{}{\mathit{vz} : \mathit{Var}\ (\Gamma, \sigma)\ \sigma}\ \emptyset \qquad\qquad \frac{x : \mathit{Var}\ \Gamma\ \tau}{\mathit{vs}\ x : \mathit{Var}\ (\Gamma, \sigma)\ \tau}\ \mathrm{weak}$$

Termes

La λ-abstraction lie la dernière variable du contexte :

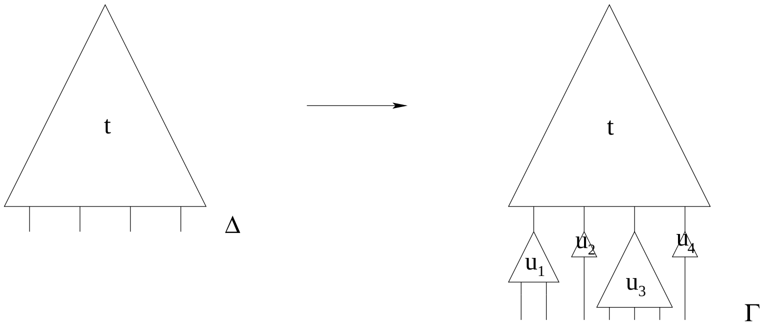
$$\frac{\Gamma : \mathit{Con} \qquad \sigma : \mathit{Ty}}{\mathit{Tm}\ \Gamma\ \sigma : \mathrm{Set}} \qquad\qquad \frac{v : \mathit{Var}\ \Gamma\ \sigma}{\mathit{var}\ \ v : \mathit{Tm}\ \Gamma\ \sigma}\ \mathrm{var}$$

$$\frac{t : \mathit{Tm}\ (\Gamma, \sigma)\ \tau}{\lambda t : \mathit{Tm}\ \Gamma\ (\sigma \to \tau)}\ \lambda \qquad \frac{t_1 : \mathit{Tm}\ \Gamma\ (\sigma \to \tau) \qquad t_2 : \mathit{Tm}\ \Gamma\ \sigma}{\mathit{app}\ \ t_1\ \ t_2 : \mathit{Tm}\ \Gamma\ \tau}\ \mathrm{app}$$

Troisième partie III

La catégorie des substitutions parallèles

Présentation



Définition

Pour tout $T : Con \to Ty \to \mathrm{Set}$:

$$\frac{\Gamma, \Delta : Con}{Subst\ \ T\ \Gamma\ \Delta : \mathrm{Set}} \qquad\qquad \frac{}{\varepsilon : Subst\ \ T\ \Gamma\ \varepsilon}\ \varepsilon$$

$$\frac{s : Subst\ \ T\ \Gamma\ \Delta \qquad t : T\ \Gamma\ \sigma}{s, t : Subst\ \ T\ \Gamma\ (\Delta, \sigma)}\ \mathrm{ext}$$

Catégorie

Prouver que les substitutions parallèles forment une catégorie est équivalent à prouver les trois théorèmes suivants :

$$\left\{ \begin{array}{rcll} id \circ s & == & s & \text{(neutralité à gauche)} \\ s \circ id & == & s & \text{(neutralité à droite)} \\ (u \circ v) \circ w & == & u \circ (v \circ w) & \text{(associativité)} \end{array} \right.$$

Produits finis (1/3)

On considère la fonction qui étend une substitution comme une fonction d'appariement :

$$\begin{array}{rcccccc} ext & : & Subst\ T\ \Gamma\ \Delta & \rightarrow & T\ \Gamma\ \sigma & \rightarrow & Subst\ T\ \Gamma\ (\Delta, \sigma) \\ & & s & \mapsto & t & \mapsto & s, t \end{array}$$

Produits finis (2/3)

Il faut alors trouver deux projecteurs :

$$\begin{cases} \pi_1 & : & \mathit{Subst}\ T\ \Gamma\ (\Delta, \sigma) \rightarrow \mathit{Subst}\ T\ \Gamma\ \Delta \\ \pi_2 & : & \mathit{Subst}\ T\ \Gamma\ (\Delta, \sigma) \rightarrow T\ \Gamma\ \sigma \end{cases}$$

vérifiant les axiomes suivants :

$$\begin{cases} \pi_1\ (\mathit{ext}\ s\ t) & == & s & \text{(first projector)} \\ \pi_2\ (\mathit{ext}\ s\ t) & == & t & \text{(second projector)} \\ \mathit{ext}\ (\pi_1\ u)\ (\pi_2\ u) & == & u & \text{(surjective pairing)} \end{cases}$$

Produits finis (3/3)

On propose :

$$\begin{cases} \pi_1 & : & s & \mapsto & s \circ id^{+\sigma} \\ \pi_2 & : & s & \mapsto & (\{var\}\ vz)[s] \end{cases}$$

Idée des preuves

- Preuves pour les variables : simples.
- Ramener les preuves sur les termes en des preuves similaires sur les variables.

Quatrième partie IV

Substitutions héréditaires et décidabilité de la $\beta\eta$-équivalence

Substitutions ponctuelles

On remplace une seule variable libre par un terme :

$$\begin{cases} (var\ x)[x := u] & = & u \\ (var\ y)[x := u] & = & var\ y \text{ si } y \neq x \\ (\lambda t)[x := u] & = & \lambda t[vs\ x := weak\ u] \\ (t_1\ t_2)[x := u] & = & t_1[x := u]\ t_2[x := u] \end{cases}$$

Règles de récriture

- β-réduction :

$$app\ (\lambda t)\ u \longrightarrow t[vz := u]$$

- η-expansion :

$$t \longrightarrow \lambda(app\ (weak\ \ t)\ (var\ \ vz))$$

Formes normales

Lemme *Le λ-calcul simplement typé muni de la β-réduction et de l'η-expansion est fortement normalisant. Les formes normales sont exactement :*

- *les λ-abstractions des formes normales ;*
- *les applications d'une variable à des formes normales de telle sorte que le résultat a un type de base (termes neutres).*

Formalisation

On peut donc définir les formes normales ainsi :

$$\frac{\Gamma : \mathit{Con} \qquad \sigma : \mathit{Ty}}{\mathit{Nf}\ \Gamma\ \sigma : \mathrm{Set}} \qquad \frac{\Gamma : \mathit{Con} \qquad \sigma, \tau : \mathit{Ty}}{\mathit{Sp}\ \Gamma\ \sigma\ \tau : \mathrm{Set}}$$

$$\frac{t : \mathit{Nf}\ (\Gamma, \sigma)\ \tau}{\lambda n\ t : \mathit{Nf}\ \Gamma\ (\sigma \to \tau)}\ \lambda\mathrm{n} \qquad \frac{}{\vec{\varepsilon} : \mathit{Sp}\ \Gamma\ \sigma\ \sigma}\ \vec{\varepsilon}$$

$$\frac{x : \mathit{Var}\ \Gamma\ \sigma \qquad \vec{ts} : \mathit{Sp}\ \Gamma\ \sigma\ \circ}{x\ \vec{ts} : \mathit{Nf}\ \Gamma\ \circ}\ \mathrm{ne} \qquad \frac{t : \mathit{Nf}\ \Gamma\ \sigma \qquad \vec{ts} : \mathit{Sp}\ \Gamma\ \tau\ \rho}{t, \vec{ts} : \mathit{Sp}\ \Gamma\ (\sigma \to \tau)\ \rho}\ \mathrm{ext}$$

$\beta\eta$-équivalence (1/2)

Lecture "dans les deux sens" des deux règles de récriture ...

$$\frac{t, u : Tm\ \Gamma\ \sigma}{t \equiv u : \mathrm{Set}} \qquad \frac{}{beta : app\ (\lambda t)\ u \equiv t[vz := u]}\ \beta$$

$$\frac{}{eta : \lambda(app\ (weak\ \ t)\ (var\ \ vz)) \equiv t}\ \eta$$

$\beta\eta$-équivalence (2/2)

... en créant une relation d'équivalence compatible avec les constructeurs des termes :

$$\frac{}{refl : t \equiv t}\ \text{refl} \qquad \frac{p : t \equiv u}{sym\ p : u \equiv t}\ \text{sym} \qquad \frac{p_1 : t \equiv u \qquad p_2 : u \equiv v}{trans\ p_1\ p_2 : t \equiv v}\ \text{trans}$$

$$\frac{p : t \equiv u}{cong\lambda\ p : \lambda t \equiv \lambda u}\ \text{cong}\lambda$$

$$\frac{p_1 : t_1 \equiv u_1 \qquad p_2 : t_2 \equiv u_2}{congApp\ p_1\ p_2 : app\ t_1\ t_2 \equiv app\ u_1\ u_2}\ \text{congApp}$$

Principe

- Substitutions du λ-calcul (simplement typé) ;
- introduites à la fin des années 1990 ;
- **structurellement récursives** ;
- fournissent un algorithme de **normalisation** dont la **terminaison se prouve par une simple récurrence** en utilisant un ordre lexicographique.

Principe : effectuer la substitution et normaliser en même temps.

$$(y\ (z, \vec{\varepsilon}))[y := \lambda n\ vz] = (\lambda n\ vz)(z, \vec{\varepsilon})$$

Notation utile

Pour pouvoir écrire :

$$t[x := u]$$

x ne doit pas être libre dans $t[x := u]$, et donc dans u ! Pour cela, on introduit une fonction permettant d'ôter une variable d'un contexte :

$$\frac{\Gamma : Con \qquad x : Var\ \Gamma\ \sigma}{\Gamma - x : Con}\ \text{min}$$

Substitution

Les règles de substitution deviennent donc :

$$\frac{t : T\ \Gamma\ \sigma \qquad x : Var\ \Gamma\ \tau \qquad u : T\ (\Gamma - x)\ \tau}{t[x := u] : T\ (\Gamma - x)\ \sigma}\ \text{subst}$$

$$\frac{\vec{ts} : Sp\ \Gamma\ \sigma\ \rho \qquad x : Var\ \Gamma\ \tau \qquad u : Nf\ (\Gamma - x)\ \tau}{\vec{ts}[x := u] : Sp\ (\Gamma - x)\ \sigma\ \rho}\ \text{substSp}$$

Substituer et normaliser en même temps

Si on ne fait que substituer, on n'obtient pas nécessairement un forme normale :

$$(y\ (z, \vec{\varepsilon}))[y := \lambda n\ vz] = (\lambda n\ vz)(z, \vec{\varepsilon})$$

Il faut normaliser en même temps, en appliquant $\lambda n\ vz$ à la chaîne $(z, \vec{\varepsilon})$. On aura donc besoin d'une troisième fonction :

$$\frac{u : Nf\ \Gamma\ \sigma \qquad \vec{ts} : Sp\ \Gamma\ \sigma\ \tau}{u \vec{@} \vec{ts} : Nf\ \Gamma\ \tau}\ \vec{@}$$

Définition

$$\text{subst} : \begin{cases} (\lambda n\ \ t)[x := u] & = & \lambda n\ \ t[x^{+vz} := u^{+vz}] \\ (x\ \ \vec{ts})[x := u] & = & u \vec{@} \vec{ts}[x := u] \\ (y\ \ \vec{ts})[x := u] & = & y^{-x}\ \ \vec{ts}[x := u] & \quad \text{si } x \neq y \end{cases}$$

$$\text{substSp} : \begin{cases} \vec{\varepsilon}[x := u] & = & \vec{\varepsilon} \\ (t, \vec{ts})[x := u] & = & t[x := u], \vec{ts}[x := u] \end{cases}$$

$$\vec{@} : \begin{cases} u \vec{@} \vec{\varepsilon} & = & u \\ (\lambda n\ \ u) \vec{@} (t, \vec{ts}) & = & (u[vz := t]) \vec{@} \vec{ts} \end{cases}$$

Théorème

Théorème *Ces définitions sont structurellement récursives.*

Idée de la preuve : considérer le cas où subst et $\vec{@}$ s'appellent mutuellement :

$$\left\{\begin{array}{rcl}(x\ (t, \vec{ts}))[x := \lambda n\ u] & = & (\lambda n\ u)\vec{@}(t', \vec{ts}') \\ & & \text{où } (t', \vec{ts}') = (t, ts)[x := u] \\ (\lambda n\ u)\vec{@}(t', \vec{ts}') & = & (u[vz := t'])\vec{@}\vec{ts}'\end{array}\right.$$

Normalisation (1/2)

Normalisons d'abord les termes neutres n'ayant pas un type de base :

$$\frac{x : Var\ \Gamma\ \sigma \qquad \vec{ts} : Sp\ \Gamma\ \sigma\ \tau}{nf'\ x\ \vec{ts} : Nf\ \Gamma\ \circ}\ \text{nf'}$$

$$\left\{\begin{array}{lcl} nf'\ x\ \vec{ts} & = & x\ \vec{ts} \\ & & \text{si } \vec{ts} : Sp\ \Gamma\ \sigma\ \circ \\ nf'\ x\ \vec{ts} & = & \lambda n\ (nf'\ x\ (append\ \vec{ts}^{+vz}\ (nf'\ vz\ \vec{\varepsilon}))) \\ & & \text{si } \vec{ts} : Sp\ \Gamma\ \sigma\ (\tau \to \rho) \end{array}\right.$$

Normalisation (2/2)

On peut maintenant normaliser tous les termes :

$$\left\{\begin{array}{rcl} nf\ (var\ v) & = & nf'\ v\ \vec{\varepsilon} \\ nf\ (\lambda t) & = & \lambda n\ (nf\ \ t) \\ nf\ (app\ t_1\ \ t_2) & = & u[vz := nf\ \ t_2] \quad \text{où } nf\ \ t_1 = \lambda n\ u \end{array}\right.$$

Opération inverse

On définit également une opération permettant de transformer une forme normale en un terme :

$$\frac{n : Nf\ \Gamma\ \sigma}{\lceil n \rceil : Tm\ \Gamma\ \sigma}\ \text{emb}$$

qui se définit simplement par pattern matching sur n.

Décidabilité

Prouver la décidabilité de la $\beta\eta$-équivalence est équivalent à prouver :

$$\frac{}{\lceil nf\ \ t \rceil \equiv t}\ \text{complétude} \qquad \frac{t \equiv u}{nf\ \ t == nf\ \ u}\ \text{consistance}$$

En effet, étant donnés deux termes t et u, il suffit de les normaliser (ce qui termine) pour obtenir t' et u'. On aura alors $t \equiv u$ si et seulement si $t' == u'$.

Stabilité

On peut également prouver la propriété suivante :

$$\frac{}{nf\ \lceil t \rceil == t}\ \text{stabilité}$$

Cinquième partie V

Conclusion

Apports et perspectives

Apports :

- bonne compréhension des substitutions parallèles ;
- une implantation des substitutions héréditaires ;
- travail utilisant un assistant de preuves.

Perspectives :

- continuer à faire des assistants de preuves ;
- développer les substitutions héréditaires (ex : normaliseur qui termine pour tous les termes du λ-calcul pur).

Plan personnel

- Approfondissement sur les substitutions parallèles ;
- découverte des substitutions héréditaires ;
- découverte d'Agda ;
- utilisation de la notation à la de Bruijn.

Un stage très complet !

Des questions ?

Huitième partie VIII

Définition des substitutions parallèles

Définition

Pour tout $T : Con \to Ty \to \mathrm{Set}$:

$$\frac{\Gamma, \Delta : Con}{Subst\ \ T\ \Gamma\ \Delta : \mathrm{Set}} \qquad\qquad \frac{}{\varepsilon : Subst\ \ T\ \Gamma\ \varepsilon}\ \varepsilon$$

$$\frac{s : Subst\ \ T\ \Gamma\ \Delta \qquad t : T\ \Gamma\ \sigma}{s, t : Subst\ \ T\ \Gamma\ (\Delta, \sigma)}\ \mathrm{ext}$$

Substitution

Pour tout $T : Con \to Ty \to \text{Set}$:

$$\frac{s : Subst\ T\ \Gamma\ \Delta \qquad t : T\ \Delta\ \sigma}{t[s] : T\ \Gamma\ \sigma}\ \text{subst}$$

Cas des variables (la substitution ne peut pas être vide) :

$$\begin{cases} vz[s,t] & = & t \\ (vs\ v)[s,t] & = & v[s] \end{cases}$$

Cas des termes ?

Affaiblissement

Extension du codomaine d'une substitution :

$$\frac{s : \mathit{Subst}\ T\ \Gamma\ \Delta \qquad \sigma : \mathit{Ty}}{s^{+\sigma} : \mathit{Subst}\ T\ (\Gamma, \sigma)\ \Delta}\ \text{weak}$$

$$\begin{cases} \varepsilon^{+\sigma} & = & \varepsilon \\ (s, t)^{+\sigma} & = & s^{+\sigma}, (\mathit{weak}\ \ t) \end{cases}$$

où :

$$\frac{t : \mathit{Tm}\ \Gamma\ \sigma}{\mathit{weak}\ \ t : \mathit{Tm}\ (\Gamma, \tau)\ \sigma}\ \text{weak}$$

Substitutions pour les termes

On peut donc maintenant définir les substitutions pour les termes :

$$\left\{ \begin{array}{rcl} (var\ \ v)[s] & = & v[s] \\ (\lambda t)[s] & = & \lambda t[s^{+\sigma}, var\ \ vz] \\ (app\ \ t_1\ \ t_2)[s] & = & app\ \ t_1[s]\ \ t_2[s] \end{array} \right.$$

s pourrait être une substitutions de variables, en modifiant juste la définition pour la λ-abstraction :

$$(\lambda t)[s] = \lambda t[s^{+\sigma}, vz]$$

Identité, composition

On a besoin d'une substitution identité et de la composition de deux substitutions :

$$\frac{}{id_\Gamma : Subst\ \Gamma\ \Gamma}\ \text{id} \qquad \frac{u : Subst\ T\ \Gamma\ \Delta \qquad v : Subst\ T\ \Delta\ \Theta}{u \circ v : Subst\ T\ \Gamma\ \Theta}\ \text{comp}$$

définies par :

$$\begin{cases} id_\varepsilon & = & \varepsilon \\ id_{\Gamma,\sigma} & = & id_\Gamma^{+\sigma}, \{var\}\ vz \end{cases} \qquad \begin{cases} u \circ \varepsilon & = & \varepsilon \\ u \circ (v, t) & = & u \circ v, t[u] \end{cases}$$