

Exercices préparatoires au théorème des développements finis et standardisation

TD du 24 novembre 2021

Alexis Saurin

Les exercices suivants étudient deux théorèmes importants du λ -calcul que l'on verra en cours : le théorème des développements finis et le théorème de standardisation.

Dans le premier cas, on étudiera une variante du λ -calcul, le $\lambda^\bullet$ -calcul, dont on montrera la forte normalisation (c'est-à-dire l'absence de suite infinie de λ -termes reliés par la β -réduction), ce qui préparera la preuve du théorème des développements finis. Dans cette variante les réductions (quelconques) permettent de représenter les développements d'ensembles de radicaux : une contrainte syntaxique imposée sur les termes permet de refléter la notion de développement dans ce sous-calcul. L'exercice suivant vise à démontrer le théorème de standardisation sous hypothèse de normalisation forte. Dans ce cas, la démonstration est grandement simplifiée et elle permet tout de même de déduire, par exemple, la standardisation du λ -calcul simplement typé que nous verrons dans le prochain cours.

Pour chaque question, on indique si c'est une question facile (★), moyenne (★★) ou un peu plus difficile (★★★). Quelques indications sont proposées à la fin du sujet.

1. Exercice préparatoire au théorème des développements finis.

Les notions nécessaires pour parler des développements finis ont été/seront vues en cours vendredi 19 novembre *mais elles ne sont pas nécessaires pour traiter l'exercice qui suit*. Elles sont toutefois utiles pour comprendre comment l'exercice s'articule avec la notion de développement (question 8) et donc sa motivation. On donne quelques définitions dans la section suivante qui ne sont pas utiles pour traiter l'exercice sur les développements finis et qui peut donc être sautée en première lecture. (ces définitions sont en revanche nécessaires pour l'exercice sur la standardisation qui vient dans la seconde section.)

1.1. Définitions préliminaires

Le théorème des développements finis permet de caractériser ce qui permet d'avoir, en λ -calcul, des séquences de réduction infinies : **la création de radicaux**. (Les **radicaux** d'un terme t sont ses sous-termes de la forme $(\lambda x.t)u$.)

Si t, u sont des λ -termes tels que $t \rightarrow_\beta u$, les radicaux de u peuvent être de deux types : (i) soient ils ont été créés par l'étape de β -réduction qui conduit de t à u , (ii) soient ils proviennent d'un radical déjà présent dans t .

Le théorème des développements finis affirme que toutes les séquences de réductions issues d'un terme t qui ne réduisent jamais un radical qui a été créé au cours de la réduction, mais uniquement des radicaux qui proviennent de radicaux présents dans le terme de départ (on parle des **résidus des radicaux** de t). Une telle séquence de réduction ne réduisant que des résidus est appelée un développement. Le théorème de développements finis affirme que tous les développements... sont finis ! (ie qu'il s'agit de séquences de β -réductions finies) et qu'elles ont même une longueur bornée en fonction du terme de départ.

La notion de **résidu** et celle de **développement** sont définies par le biais des **occurrences** qui permettent de localiser les sous-termes d'un λ -terme par le biais d'un système d'adressage dans l'arbre du λ -terme : on désigne la position d'un sous-terme dans un λ -terme par un mot sur l'alphabet $\{0, 1, 2\}$:

Définition .1 (Ensemble des occurrences d'un λ -terme)

$\text{Occ}(t)$ est un ensemble de mots (ou de suites finies) sur l'alphabet $\{0, 1, 2\}$ défini comme suit :

- $\text{Occ}(x) = \{\epsilon\}$;
- $\text{Occ}((t)u) = \{\epsilon\} \cup 1 \cdot \text{Occ}(t) \cup 2 \cdot \text{Occ}(u)$;
- $\text{Occ}(\lambda x.t) = \{\epsilon\} \cup 0 \cdot \text{Occ}(t)$

On notera $a, b, c, d, a', b', c', \dots$ les occurrences.

Si t est un λ -terme et $a \in \text{Occ}(t)$, on peut alors définir le **sous-terme de t à l'occurrence a** , $t|_a$ comme suit :

$$\begin{array}{ll} t|_\epsilon = t & C_{(t,\epsilon)} = \square \\ \lambda x.t|_{0 \cdot a} = t|_a & C_{(\lambda x.t, 0 \cdot a)} = \lambda x.C_{(t,a)} \\ (t)u|_{1 \cdot a} = t|_a & C_{((t)u, 1 \cdot a)} = (C_{(t,a)})u \\ (t)u|_{2 \cdot a} = u|_a & C_{((t)u, 2 \cdot a)} = (t)C_{(u,a)} \end{array}$$

On peut définir l'ensemble des occurrences de t où se trouvent des radicaux de t :

Définition .2 (Radicaux)

On notera $\text{RAD}(t)$ l'ensemble des occurrences des radicaux de t . (on rappelle qu'un radical est une expression de la forme $(\lambda x.t)u$.)

Si t, u sont des λ -termes tels que $t \rightarrow_\beta u$, c'est nécessairement que $\text{RAD}(t) \neq \emptyset$: chaque occurrence de radical est la position d'une possible étape de β -réduction. On parlera ainsi de réduction à l'occurrence a :

Définition .3 (Réduction à l'occurrence a)

On écrira $t \xrightarrow{a} u$ si t se réduit en u en réduisant le radical d'occurrence a .

On peut alors raffiner la notion de séquence de β -réduction vue au début du cours de λ -calcul en n'intégrant pas uniquement les termes reliés par $\rightarrow_\beta$, mais également les occurrences où se passent la réduction :

Définition .4 (Séquence de réduction)

Une réduction est une suite $\sigma = (t_i, a_i)_{i \in \lambda \in \omega+1}$ (finie ou infinie) d'élément de $\Lambda \times \{0, 1, 2\}^*$ tel que pour tout $i \in \lambda$, $a_i \in \text{RAD}(t_i)$ et tel que si $i+1 \in \lambda$, $t_i \xrightarrow{a_i} t_{i+1}$.

Si t, u sont des λ -termes tels que $t \rightarrow_\beta u$, les radicaux de u peuvent être de deux types : soient ils ont été créés par l'étape de β -réduction qui conduit de t à u , soient ils proviennent d'un radical déjà présent dans t . En effet, les radicaux de t , lors d'une étape de β -réduction en a , peuvent être soit (i) consommés (s'il s'agit du radical en a), (ii) soit être déplacés, (iii) soit être effacés, (iv) soit être substitués ailleurs dans le terme, une ou plusieurs fois. Par ailleurs l'opération de β -réduction peut créer de nouveaux radicaux comme vu en cours.

Informellement, les **résidus** d'un radical de t après une réduction $t \xrightarrow{\sigma} u$ sont l'ensemble des radicaux de u qui proviennent de ce radical. Si $t \in \Lambda$, $a, b \in \text{RAD}(t)$ et si u est le terme obtenu en réduisant le radical d'occurrence b de t ($t \xrightarrow{b} u$), on désigne par a/b l'ensemble des résidus du radical a de t dans u après l'étape de réduction $t \xrightarrow{b} u$. Le détail de cette notion est traitée en cours.

L'intuition qu'il faut avoir en tête est que dans un **développement** d'un ensemble de radicaux, on ne réduit que les radicaux contenus dans cet ensemble (et leurs résidus) et on s'interdit de réduire d'autres radicaux du terme initial de même que les radicaux créés au cours de la réduction. Un développement σ de $\mathcal{F}$ est dit **complet** si l'ensemble des résidus de $\mathcal{F}$ après σ est vide. On peut maintenant énoncer le théorème des développements finis :

Théorème 1 (des développements finis)

Soient $t \in \Lambda$, $\mathcal{F} \subseteq \text{RAD}(t)$ (ensemble de radicaux de t), $a \in \text{RAD}(t)$ et σ et τ deux développements complets de $\mathcal{F}$; alors :

1. tous les développements relatifs à $\mathcal{F}$ sont finis et de longueur bornée ;
2. σ et τ aboutissent au même terme (si $t \xrightarrow{\sigma} u$ et $t \xrightarrow{\tau} v$ alors $u \equiv v$) et
3. σ et τ induisent les mêmes résidus (si $a \in \text{RAD}(t)$, alors $a/\sigma = a/\tau$).

et passer à l'exercice :

1.2. Normalisation forte du $\lambda^\bullet$ -calcul

Le but de l'exercice est d'étudier la normalisation forte (SN) d'une restriction du λ -calcul, que l'on appellera le $\lambda^\bullet$ -calcul. Il s'agit donc de démontrer le théorème suivant :

Théorème 2

$\lambda^\bullet$ est fortement normalisant (c'est-à-dire qu'il n'existe pas de suite infinie de β -réductions dans ce calcul).
On définit le $\lambda^\bullet$ -calcul :

Définition 3 ($\lambda^\bullet$ -calcul)

On ajoute au λ -calcul une nouvelle variable que l'on notera $\bullet$ et on définit un nouvel ensemble de termes $\Lambda^\bullet$ par la grammaire suivante (x désignant une variable quelconque du λ -calcul habituel, ie $x \neq \bullet$) :

$$\Lambda^\bullet \ni t, u, v := x \mid \lambda x.t \mid (\bullet)tu \mid (\lambda x.t)u$$

C'est-à-dire que $\Lambda^\bullet$ est le plus petit ensemble tel que :

- Si $x \neq \bullet$, $x \in \Lambda^\bullet$;
- Si $x \neq \bullet$ et $t \in \Lambda^\bullet$, alors $\lambda x.t \in \Lambda^\bullet$;
- Si $t, u \in \Lambda^\bullet$, alors $(\bullet)tu \in \Lambda^\bullet$;
- Si $t, u \in \Lambda^\bullet$ et si t est une λ -abstraction, alors $(t)u \in \Lambda^\bullet$.

Question 1 (★)

Quelle est la particularité de $\bullet$ dans le $\lambda^\bullet$ -calcul ?

Expliciter, sur la notation arborescente des $\lambda^\bullet$ -termes, les situations où le nœud application peut apparaître.

L'idée de la nouvelle variable $\bullet$ est de “geler” certains radicaux et d'empêcher les créations de radicaux.¹

Question 2 (★★)

Montrer que si $t, u \in \Lambda^\bullet$ et $x \neq \bullet$, alors $u\{t/x\} \in \Lambda^\bullet$.

Question 3 (★)

En déduire que si $t \in \Lambda^\bullet$ et $t \rightarrow_\beta^* u$, alors $u \in \Lambda^\bullet$.

La question précédente nous assure que $\Lambda^\bullet$ est donc stable par β -réduction. La suite de l'exercice vise à montrer que $\Lambda^\bullet$ est fortement normalisant.

Question 4 (★★)

Expliquer pourquoi une simple induction sur la structure du terme t ne suffit pas à montrer qu'il est SN.

On va en fait surcharger l'hypothèse d'induction en introduisant la notion de substitution fortement normalisante. À côté de la substitution habituelle, $t\{u/x\}$, on introduit une nouvelle notion de substitution «en parallèle» :

Définition 4 (Substitution parallèle)

Une **substitution parallèle** est une liste $[x_1 := t_1, \dots, x_n := t_n]$ où les variables $x_1, \dots, x_n$ sont deux à deux disjointes et toutes différentes de $\bullet$. L'application de la substitution parallèle à t est définie par :

- $x_i[\sigma] = t_i$ ($1 \leq i \leq n$),
- $y[\sigma] = y$ pour toute variable hors de $\{x_1, \dots, x_n\}$,
- $((s)t)[\sigma] = (s[\sigma])t[\sigma]$, et
- $(\lambda x.t)[\sigma] = \lambda x.(t)[\sigma]$, où x n'est libre dans aucun t_i , et est différent de x_i , pour tout i , $1 \leq i \leq n$.

La substitution usuelle est une instance de la substitution parallèle, mais on notera qu'en général, par exemple, $t[x_1 := t_1, x_2 := t_2] \neq t\{t_1/x_1\}\{t_2/x_2\}$.

Définition 5 (Substitution fortement normalisante)

Une substitution parallèle σ est dite **fortement normalisante** (SN) si tous les t_i le sont, $1 \leq i \leq n$.

Question 5 (★)

Montrer que pour une certaine substitution parallèle $[\sigma]$, t est SN si, et seulement si, $t[\sigma]$ est SN.

Question 6 (★★★)

Montrer que $t[\sigma]$ est SN pour tout $t \in \Lambda^\bullet$ et toute substitution parallèle σ SN. En déduire que $\Lambda^\bullet$ est SN.

1. Dans le problème, à part pour la dernière question, on aurait aussi pu travailler dans un calcul où $\bullet$ était en fait une constante sans δ -règle associée ; cela ne changerait rien de fondamental aux preuves.

Question 7 (★★)

Expliquer pourquoi la considération de $t[\sigma]$ a permis de dépasser le problème identifié à la question 4.

Question 8 (★★★ – question d’ouverture)

Proposer une manière d’associer, à tout $\lambda^\bullet$ -terme T , un λ -terme t qui n’utilise pas la variable $\bullet$. Si l’on voit T comme un λ -terme (c’est-à-dire qu’on cesse de traiter $\bullet$ de manière spécifique), il y a une manière de réaliser cette transformation de manière interne au calcul ; l’expliciter. Comparez les réductions possibles à partir de T et celles à partir de t . Comment relier la normalisation forte du $\lambda^\bullet$ -calcul au théorème des développements finis ?

2. Standardisation sous hypothèse de normalisation forte

On rappelle que dans un λ -terme qui n’est pas en forme normale, on peut toujours trouver un radical plus à gauche que tous les autres : la réduction gauche consiste à réduire toujours le radical le plus à gauche.

(On dit qu’un radical a est à gauche d’un radical b dans t si la position du motif (λ de a est à gauche du motif (λ de b dans t))

On notera $t \rightarrow_g u$ le fait que t se réduit en u par une étape de réduction gauche.

Définition 6 (Réduction standard)

Une réduction $\sigma = (t_i, a_i)_{i \in \lambda \in \omega+1}$ (finie ou infinie) est dite **standard** si pour tous i, j tels que $i < j \in I$, il n’y a pas de radical $b \in \text{RAD}(t_i)$ à gauche de a_i tel que $a_j \in b/(t_k, a_k)_{i \leq k < j}$.

S’il existe une dérivation standard σ , $t \xrightarrow{\sigma} u$, on note $t \xrightarrow{st} u$.

Théorème 7 (Standardisation dans le cas SN)

Soient $t, u \in \Lambda$, t fortement normalisant. Si $t \xrightarrow{\beta}^* u$ alors $t \xrightarrow{st} u$.

Définition 8 (Val(t))

Soit $t \in \Lambda$. Si $t \xrightarrow{\beta}^* \lambda x.u$, on note $\text{Val}(t)$ le premier terme commençant par une abstraction qui apparaît dans la réduction.

Question 9 (★)

Pourquoi la définition précédente est-elle bien définie ?

Question 10 (★★)

Soient $t, u \in \Lambda$ et σ une réduction standard de t à $\lambda x.u$.

Montrer qu’il existe des réductions ρ et τ telles que $\sigma = \rho \cdot \tau$ où $t \xrightarrow{\rho} \text{Val}(t)$, $\text{Val}(t) \xrightarrow{\tau} u$, ρ est une réduction de tête et τ est standard.

Question 11 (★★)

Montrer que si $t, u \in \Lambda$, t fortement normalisant, sont tels que $t \xrightarrow{\beta}^* u$, alors $t \xrightarrow{st} u$.

On procèdera par induction ; la mesure n’est pas immédiate à trouver et une indication est donnée ci-dessous.

Indications

1. Pas d’indication particulière.
2. Choisir le bon terme pour faire une induction structurelle.
3. Traiter le cas d’une réduction de longueur 1, le reste suit facilement.
4. Tenter un raisonnement par induction sur la structure de t et voir où ça bloque.
5. Essayer avec la substitution parallèle la plus simple qui soit.
6. Il s’agit de la question longue de l’exercice.
 - a) Ce coup-ci, on fait une induction structurelle sur t , cela va marcher.
 - b) Voyez-vous une autre manière d’écrire $(t(u/x))[\sigma]$ qui mette à profit la substitution parallèle ?
7. Pensez à l’information que vous pouvez conserver en passant de $(\lambda x.t)u[\sigma]$ à $t[x := u[\sigma], \sigma]$.
8. On ne donne une indication que sur le début de la question, le reste sera l’objet du cours...
 - a) La transformation ne consiste pas à choisir une variable x et à poser $t = T\{x/\bullet\}$...
 - b) Si $\bullet$ est une variable comme une autre, elle peut être affectée par une substitution. Par quel terme pourrait-on substituer $\bullet$ dans T pour obtenir t ?
9. Pas d’indication particulière.
10. Pas d’indication particulière.
11. Ni la structure de t , ni la longueur de la plus longue réduction issue de t ne suffisent, à elles seules, pour mener à bien l’induction. Tentez de combiner les deux.