

M2 LMFI – Cours Fondamental de logique

THÉORIE DE LA DÉMONSTRATION

λ-calcul simplement typé et théorèmes de normalisation

(version préliminaire)

Alexis Saurin

Cours du 3 décembre 2021

Table des matières

1	Définition du λ-calcul simplement typé, $\Lambda^{\rightarrow}$	1
1.1	$\Lambda^{\rightarrow}$ à la Church	2
1.2	$\Lambda^{\rightarrow}$ à la Curry	2
1.3	Relations entre les syntaxes à la Church et à la Curry	2
1.4	Quelques exemples de termes typés	3
1.5	Réductions dans les λ-calculs simplement typés	5
2	Théorème de normalisation faible de $\Lambda^{\rightarrow}$	6
2.1	Neutralité, réductibilité	7
2.2	Énoncé du théorème et structure de la preuve	8
2.3	Preuve des deux lemmes clés	8
2.4	Vers la normalisation forte	9
3	Théorème de normalisation forte de $\Lambda^{\rightarrow}$	9
3.1	Neutralité, réductibilité	9
3.2	Énoncé du théorème et structure de la preuve	10
3.3	Preuve des deux lemmes clés	11

1 Définition du λ-calcul simplement typé, $\Lambda^{\rightarrow}$

On introduit dans cette section deux variantes du lambda-calcul simplement typé, selon les styles à la Church et à la Curry respectivement, que l'on relie ensuite.

Définition 1.1 (*Types simples*)

On se donne un ensemble infini dénombrable $\mathcal{T}_{\text{At}}$ dit des **types atomiques**. Les types simples sont définis inductivement par la grammaire :

$$T, U, V ::= A \mid U \rightarrow V \quad A \in \mathcal{T}_{\text{At}}.$$

1.1 $\Lambda^\rightarrow$ à la Church

Le λ -calcul simplement typé à la Church est l'ensemble défini inductivement (conjointement avec leur type) comme suit :

Définition 1.2 ($\Lambda^\rightarrow$ à la Church)

On considère, pour chaque type simple T , un ensemble infini dénombrable de variables de ce type, $\mathcal{V}^T$, ces ensembles étant disjoints deux-à-deux.

Les termes simplement typés à la Church sont le plus petit ensemble tel que :

- Pour toute variable x de $\mathcal{V}^T$, x^T est un terme de type T (de variable libre x) : $x^T : T$;
- Pour tout terme v de type V et toute toute variable x de $\mathcal{V}^U$, $\lambda x^U.v$ est un terme de type $U \rightarrow V$ (de variables libres $fv(v) \setminus \{x\}$) : $\lambda x^U.v : U \rightarrow V$;
- Pour tous termes t et u de types respectifs $U \rightarrow T$ et U , $(t)u$ est un terme de type T (de variables libres $fv(t) \cup fv(u)$) : $(t)u : T$;

ce qu'on peut résumer par le système d'inférence suivant :

$$\frac{}{x^U : U} (Var) \quad (x \in \mathcal{V}^U) \quad \frac{t : T}{\lambda x^U.t : U \rightarrow T} (Abs) \quad (x \in \mathcal{V}^U) \quad \frac{t : U \rightarrow T \quad u : U}{(t)u : T} (App)$$

On remarque que pour tout λ -terme à la Church t , il existe un type T et un seul, tel que $t : T$.

1.2 $\Lambda^\rightarrow$ à la Curry

Il s'agit du calcul sur les λ -termes purs associé à une notion de *typabilité*.

Cette relation de typabilité met en relation un **contexte de typage** Γ , un **λ -terme** t et un **type** T , et est définie au moyen d'un système de dérivation dont les règles d'inférence sont inspirées de la déduction naturelle implicative minimale, c'est-à-dire qu'il s'agit du lambda-calcul (pur) muni du système de type suivant trois règles de typage données plus bas.

Les contextes de typage sont des fonctions à support fini de l'ensemble des variables dans l'ensemble des types simples, représentées syntaxiquement sous la forme $\{x_1 : T_1, \dots, x_n : T_n\}$ ou encore définies par la grammaire :

$$\Gamma ::= \cdot \mid \Gamma, x : T,$$

en exigeant qu'une variable apparaisse au plus une fois dans un contexte.

$\frac{}{\Gamma, x : T \vdash x : T} (Var) \quad \frac{\Gamma, x : U \vdash t : T}{\Gamma \vdash \lambda x.t : U \rightarrow T} (Abs), \quad x \notin \Gamma \quad \frac{\Gamma \vdash t : U \rightarrow T \quad \Gamma \vdash u : U}{\Gamma \vdash (t)u : T} (App)$
--

On notera le fait qu'un λ -terme t apparaît en conclusion d'une dérivation de typage dans le contexte Γ et avec le type T par le jugement :

$$\Gamma \vdash_{\text{STLC}} t : T.$$

1.3 Relations entre les syntaxes à la Church et à la Curry

On a donc deux approches du calcul simplement typé :

- l'une dite à la Church, où les termes portent des indications de type qui les mettent en correspondance avec les déductions de la logique intuitionniste (ou plutôt minimale) implicative, dans le style des déductions comme arbre de formules.
- l'autre dite à la Curry, dans laquelle on définit une relation de typage, ternaire, $\vdash_{\text{STLC}}$ qui met en relation un contexte de typage Γ , un λ -terme pur (non typé) t et un type simple, T , ce qu'on note $\Gamma \vdash_{\text{STLC}} t : T$. Ici, ce sont les dérivations de typage qui sont en relation avec la logique minimale, dans sa présentation à base de séquents.

Remarque 1.3

Dans les deux cas, le rapport avec la déduction naturelle se fait une par une gestion différente des hypothèses non déchargées et du déchargement des hypothèses – qui est ici géré par la notion de liaison de variables. La correspondance est exacte pour les termes clos.

On va maintenant relier les deux présentations des types simples au moyen d’une fonction d’oubli qui enverra les termes à la Church dans les termes typables à la Curry et dont on étudiera les propriétés par rapport à la normalisation. Cela nous permettra de montrer l’équivalence entre normalisation faible (resp. forte) dans le calcul à la Church et normalisation faible (resp. forte) dans le calcul à la Curry.

Définition 1.4 (fonction d’oubli)

On définit inductivement une **fonction d’oubli** des types des λ -termes à la Church dans les λ -termes purs :

- $[x^T]^- = x$;
- $[\lambda x^T. t]^- = \lambda x. [t]^-$;
- $[(t) u]^- = ([t]^-) [u]^-$.

On remarque que l’oubli est compatible avec la substitution : $[t \{u/x^U\}]^- = [t]^- \{[u]^-/x\}$, pour tous t, u et x dont les types sont compatibles (ie $u : U$). Ceci vient du fait que, dans le calcul à la Church, les ensembles de variables des différents types sont deux à deux disjoints et donc que $[x^T]^- \neq [y^U]^-$ dès que $T \neq U$.

On peut alors montrer que l’oubli envoie bien les termes de $\Lambda^\rightarrow$ à la Church dans des termes purs typables dans le calcul à la Curry, du type attendu :

Proposition 1.5

Soit $t : T$ un terme à la Church de $\Lambda^\rightarrow$ dont les variables libres sont données par la famille $(x_i^{T_i})_{1 \leq i \leq n}$. Alors $x_1 : T_1, \dots, x_n : T_n \vdash_{\text{STLC}} [t]^- : T$ est dérivable dans $\Lambda^\rightarrow$.

Démonstration : Le résultat précédent se prouve par une simple induction sur la structure du terme t . □

Il est évident que, réciproquement :

Proposition 1.6

Si δ est une dérivation de type d’un jugement $\Gamma \vdash_{\text{STLC}} t : T$, telle que pour tout $x_i : T_i \in \Gamma$, $x_i \in \mathcal{V}^{T_i}$, alors δ est isomorphe à un terme $u : T$ de $\Lambda^\rightarrow$ à la Church dont les variables libres sont parmi les variables de Γ . En outre, $[u]^-$ est α -équivalent à t .

Remarque 1.7

Le seul point légèrement subtil concerne les variables et l’ α -équivalence que l’on n’a pas défini dans le calcul à la Church. L’ α -équivalence doit ici se faire au sein de l’ensemble des variables d’un type donné (de manière similaire au cas de la quantification dans les logiques multisortées).

Dans la suite, on considèrera les deux versions de λ -calcul simplement typés modulo α -équivalence.

1.4 Quelques exemples de termes typés

Dans cette section, on propose quelques exercices pour se familiariser avec les termes types en reprenant des termes vu au fil du cours et en se demandant s’ils sont, ou non, typables à la Curry, et le cas échéant en fournissant le(s) type(s) qui leur sont associés.

On supposera dans la suite qu’on dispose d’un type atomique $o \in \mathcal{T}_{\text{At}}$.

Exercice 1.1

Trouver un type T , le plus simple possible, et donner une dérivation de typage pour l’identité, $\lambda x. x$.

Donner le terme typé à la Church correspondant.

Exercice 1.2

On rappelle l'encodage de Church des paires :

- $\text{paire} = \lambda x. \lambda y. \lambda p. (p)xy$
- $\pi_1 = \lambda y. (y) \lambda x_1. \lambda x_2. x_1$
- $\pi_2 = \lambda y. (y) \lambda x_1. \lambda x_2. x_2$

Par souci de lisibilité, on notera $\langle t, u \rangle$ le terme $(\text{paire})tu$.

- Montrer que si A, B, C sont des types, on a $\Gamma \vdash_{\text{STLC}} \text{paire} : A \rightarrow (B \rightarrow (A \rightarrow (B \rightarrow C)) \rightarrow C)$.
- Donner une dérivation de typage pour chacun des deux autres termes π_1 et π_2 . Trouve-t-on les types auxquels on s'attendrait ?
- Donner pour chaque dérivation de typage ci-dessus, le terme typé à la Church correspondant.

Exercice 1.3

Peut-on typer le terme $\Delta = \lambda x. (x)x$ dans le λ -calcul simplement typé ?

Exercice 1.4

On note Nat le type $(o \rightarrow o) \rightarrow (o \rightarrow o)$. Montrer que pour tout $n \in \mathbb{N}$, $\Gamma \vdash_{\text{STLC}} \bar{n} : \text{Nat}$.

- Donner l'ensemble de tous les types T tels que pour tout $n \in \mathbb{N}$, $\Gamma \vdash_{\text{STLC}} \bar{n} : T$.
- Donner pour chaque dérivation de typage obtenue, le terme typé à la Church correspondant.
- Caractériser tous les termes en forme normale qui sont typables de type Nat .

Exercice 1.5

On rappelle la définition des encodages de fonctions arithmétiques de base :

- $\text{Succ} = \lambda n. \lambda f. \lambda x. (f)(n)fx$;
- $\text{Add} = \lambda m. \lambda n. \lambda f. \lambda x. (m)f(n)fx$;
- $\text{Mult} = \lambda m. \lambda n. \lambda f. \lambda x. ((m)(n)fx)$;
- $\text{Exp} = \lambda m. \lambda n. \lambda f. \lambda x. (m)nfx$;
- $\text{Pred} = \lambda n. (\pi_1)((n)\lambda p. (\lambda p_2. \langle p_2, (\text{Succ})p_2 \rangle)(\pi_2)p) \langle \bar{0}, \bar{0} \rangle$.

Montrer que

- $\Gamma \vdash_{\text{STLC}} \text{Succ} : \text{Nat} \rightarrow \text{Nat}$.
- $\Gamma \vdash_{\text{STLC}} \text{Add} : \text{Nat} \rightarrow (\text{Nat} \rightarrow \text{Nat})$.
- $\Gamma \vdash_{\text{STLC}} \text{Mult} : \text{Nat} \rightarrow (\text{Nat} \rightarrow \text{Nat})$.
- Exp est typable mais qu'elle n'est pas typable du type attendu $\text{Nat} \rightarrow (\text{Nat} \rightarrow \text{Nat})$.
- Que dire de Pred ?

Donner pour chaque dérivation de typage ci-dessus, le terme typé à la Church correspondant.

Exercice 1.6

On rappelle que le test à zéro est représenté par :

$$Z? = \lambda xyz. ((z)\lambda x'. y)x$$

et on rappelle que ce terme a la propriété suivante :

$$\begin{array}{ll} (Z?)xy\bar{0} & \longrightarrow^* x \\ (Z?)xy\bar{n+1} & \longrightarrow^* y \\ (Z?)xyv & \text{n'est pas résoluble si } v \text{ n'est pas résoluble.} \end{array}$$

- Soit T un type. Pour quel type U typant les entiers de Church (voir la question 1.4), a-t-on : $\Gamma \vdash_{\text{STLC}} Z? : T \rightarrow (T \rightarrow (U \rightarrow T))$.
- Donner, le terme typé à la Church correspondant à la dérivation de typage de $Z?$.

Exercice 1.7 (*À propos de booléens et de test conditionnel*)

On note Bool le type $o \rightarrow (o \rightarrow o)$.

- Combien y a-t-il de termes en forme normale qui sont typables de ce type ?
- On note maintenant Bool^T le type $T \rightarrow (T \rightarrow T)$ pour T un type quelconque. Donner un λ -terme IfThenElse typable à la Curry de type $\text{Bool}^T \rightarrow \text{Bool}^T$ et deux termes $\text{true}, \text{false}$ typables de type Bool^T qui permettent de représenter le test conditionnel “retournant” dans le type T , ie tels que si u, v sont typables de type T ,

$$(\text{IfThenElse})\text{true } u \ v \longrightarrow^* u \quad \text{et} \quad (\text{IfThenElse})\text{false } u \ v \longrightarrow^* v$$

- Donner pour chaque dérivation de typage ci-dessus, le terme typé à la Church correspondant.
- Comparer avec l'exercice 1.2.

1.5 Réductions dans les λ -calculs simplement typés

Plaçons-nous dans le système à la Curry des λ -termes typables.

Si $\Gamma \vdash_{\text{STLC}} t : T$ et $t \longrightarrow_{\beta} u$, il est naturel de se demander comment se comporte u vis-à-vis des types simples.

Les variables libres de u , étant contenues dans les variables libres de t , la question de la typabilité de u dans le contexte Γ fait sens et on a en effet :

Proposition 1.8 (*Préservation du typage, aussi dite réduction du sujet, version à la Curry*)

Soient $t, u \in \Lambda$, T un type simple et Γ un contexte de typage couvrant les variables libres de t . Alors si $\Gamma \vdash_{\text{STLC}} t : T$ et $t \longrightarrow_{\beta} u$, on a $\Gamma \vdash_{\text{STLC}} u : T$.

La propriété ci-dessus et les relations précédemment établies entre termes à la Church et à la Curry suggèrent que la β -réduction fait sens aussi dans le calcul typé à la Church. En effet :

Proposition 1.9

Si T, U sont des types simples, $t : T$ et $u : U$ sont des termes à la Church et x^U est une variable, alors $t\{u/x\}$ est un terme de $\Lambda^{\rightarrow}$ à la Church et $t\{u/x\} : T$.

Définition 1.10 (β -réduction typée)

La β -réduction typée est la plus petite relation compatible sur les termes typés contenant : $(\lambda x^U. t)u \longrightarrow_{\beta} t\{u/x\}$, pour T, U des types simples, $t : T$ et $u : U$ sont des termes à la Church et x^U est une variable.

Les réductions dans le système à la Church et dans le système à la Curry peuvent également être comparées :

Proposition 1.11

Soient t, u des termes à la Church.

1. t se réduit sur u par une étape de β si, et seulement si $[t]^-$ se réduit sur $[u]^-$ par une étape de β .
2. t est en forme normale si, et seulement si, $[t]^-$ est en forme normale.

Démonstration : Le premier point se démontre par induction sur la structure de t : les cas de la variable et de la λ -abstraction ne posent pas de difficulté. Pour l'application, $t = (v)w$, on raisonne comme suit :

- $\Rightarrow$ est évident.
- $\Leftarrow$ dans le sens réciproque, si le radical réduit se trouve dans $[v]^-$ ou $[w]^-$, on a recours à l'hypothèse d'induction pour conclure. Sinon, $[t]^-$ est lui-même le radical et c'est-à-dire $[v]^- = \lambda x. [v']^-$ pour un certain v' tel que $v = \lambda x. v'$ et $u = [v']^- \{ [w]^- / x \}$, ce qui permet de conclure.

Le second point découle immédiatement du premier. □

Finalement, on obtient le résultat qui nous intéresse :

Théorème 1.12

La normalisation faible (resp. forte) de $\Lambda^{\rightarrow}$ à la Church est équivalente à la normalisation faible (resp. forte) de $\Lambda^{\rightarrow}$ à la Curry.

Démonstration : Pour la normalisation faible, le raisonnement est direct et très simple : supposons $\Lambda^{\rightarrow}$ à la Church faiblement normalisant et soit t typable dans le système à la Curry. On sait qu'il existe u tel que $[u]^{-} = t$ et que par hypothèse, u a une forme normale v . La proposition précédente nous assure que $t \longrightarrow^{*} [v]^{-}$ qui est normal. Dans l'autre direction, si tout terme typable dans $\Lambda^{\rightarrow}$ à la Curry normalise, considérons un terme à la Church, t . On sait alors que $[t]^{-}$ normalise vers u et qu'il existe v un terme à la Church tel que $t \longrightarrow^{*} v$ et $[v]^{-} = u$. Par la proposition, on a sait que v est normal d'où le résultat.

Pour la normalisation forte, le raisonnement est tout aussi direct. Supposons qu'il existe un terme t typable à la Curry dont est issue une séquence de réduction infinie. On sait qu'il existe un terme à la Church dont l'oubli est égal à t et les remarques précédentes nous assurent qu'une réduction infinie est issue de ce terme. Par contraposition, on voit que la normalisation forte du calcul à la Church assure que tout terme typable à la Curry normalise fortement. Réciproquement, supposons maintenant qu'il existe un terme t à la Church dont est issue une réduction infinie : la remarque précédente nous assure que cette réduction contient une infinité d'étapes de β -réductions d'où l'on déduit que de $[t]^{-}$ est issue une réduction infinie. □

La préservation du typage pour $\Lambda^{\rightarrow}$ à la Curry nous dit en fait que le lambda-calcul pur est une extension conservative du lambda-calcul à la Curry et permet notamment d'importer bon nombre de résultats du calcul pur sans avoir à les redémontrer. On a ainsi :

- Confluence ;
- Unicité de la forme normale ;
- Normalisation gauche.

(Les résultats d'existence, en revanche ne peuvent évidemment pas nécessairement s'importer... et c'est heureux puisqu'on veut justement éviter qu'il puisse exister un terme qui aient des réductions infinies !)

En plus de cela, on va voir que le calcul typé se comporte mieux que le calcul pur vis-à-vis de la normalisation. Dans les deux prochaines sections, on verra les théorèmes de normalisation faible, puis de normalisation forte, du λ -calcul simplement typé.

Pour bien clarifier les idées, on rappelle que :

- un λ -terme est **faiblement normalisant** s'il possède une réduction aboutissant à une forme normale.
- un λ -terme est **fortement normalisant** s'il ne possède pas de réduction infinie, c'est-à-dire si quelque soit le choix du radical qu'on réduit à chaque étape, on finit par aboutir à une forme normale.
- On dira qu'un calcul est faiblement (resp. fortement) normalisant si tous ses termes sont faiblement (resp. fortement) normalisants.

2 Théorème de normalisation faible de $\Lambda^{\rightarrow}$

Dans cette section on démontre la normalisation faible du λ -calcul simplement typé. On va utiliser une technique de preuve dite *par réductibilité*, technique introduite par Tait en 1967.

Pour donner une intuition rapide de la difficulté du problème :

- On cherche à démontrer le résultat de manière inductive en trouvant donc la bonne mesure bien fondée qui décroisse avec la réduction. Une fois le résultat établi, on sait qu'une telle mesure existe : il suffit de prendre le plus court chemin menant à une forme normale. Bien sûr, c'est impossible pour construire la preuve...

- Une mesure candidate est la complexité du terme (en faisant une induction sur la structure du terme). Cette mesure n'est pas assez fine puisque lorsqu'on effectue une étape de β -réduction, on peut accroître la complexité d'un terme notamment du fait des duplications.
- Une autre mesure candidate est la complexité du type, mais à nouveau, cette mesure est trop simpliste puisque la normalisation (faible ou forte) d'un terme dépend évidemment de la structure de ses sous-termes et que ceux-ci peuvent être de type arbitrairement plus compliqués que le terme lui-même, comme par exemple avec $(\lambda x^U.t)u : T$ avec U un type qui arbitrairement plus compliqué que T .
- La solution va résider dans le fait d'articuler les structures inductives des termes et des types. C'est ce que fait la construction par réductibilité... même si c'est un peu caché.

Remarquons qu'on a déjà vu, dans le cadre de la déduction naturelle, une preuve de normalisation faible que l'on peut directement appliquée au λ -calcul simplement typé, consistant à construire une mesure explicite : on appelle **type d'un radical** $(\lambda x^U.t)u : T$ le type $U \rightarrow T$ et on définit le poids d'un terme t , $w(t)$, comme étant la paire de la taille maximale du type d'un radical de t et du nombre de radical de cette taille maximale. Étant donné un terme t , on remarque facilement que si on réduit, dans t , un radical dont le type est de taille maximale et tel que le radical ne contient pas de radical de même taille, alors on obtient un réduit t' de t tel que $w(t') <_{lex} w(t)$, ce qui garantit que l'on atteint ainsi une forme normale.

Pourquoi, alors, chercher une autre démonstration ? Parce que la méthode de preuve par réductibilité est une technique très générale : (i) qui s'étend à la normalisation forte du λ -calcul simplement typé et (ii) qui se généralisera à des cadres plus riches, comme le système F pour lesquels on ne dispose pas nécessairement de preuve construite sur une mesure inductive explicite.

On se place dans le λ -calcul simplement typé à la Church, où les termes sont porteurs d'annotations de typage. On donne ici une preuve de la normalisation faible du λ -calcul simplement typé (à la Church) que l'on généralisera plus tard à une preuve de normalisation forte.

2.1 Neutralité, réductibilité

On commence par identifier, pour chaque type T , deux ensembles de termes de ce type.

Définition 2.1 (*Terme neutre*)

On appellera **neutre** un λ -terme qui ne débute pas par une abstraction.

La propriété des termes neutres qui nous intéressera dans la suite est que si t est un terme neutre, alors pour tout terme u , les réduits en une étape de $(t)u$ sont soit de la forme $(t')u$ avec t' réduit de t soit de la forme $(t)u'$ avec u' réduit de $u : t$ et u ne peuvent pas interagir lors de la première étape d'une telle réduction.

On définit alors $\text{Neut}(T)$, $\text{Norm}(T)$ de la manière suivante :

Définition 2.2 ($\text{Neut}(T)$, $\text{Norm}(T)$)

$$\text{Neut}(T) = \{t \in \Lambda^{\rightarrow}; t \text{ normal et neutre de type } T\}$$

$$\text{Norm}(T) = \{t \in \Lambda^{\rightarrow}; t \text{ normalisable de type } T\}.$$

Par ailleurs, à chaque type T on associe un ensemble de termes réductibles de type T :

Définition 2.3 ($\text{RED}^{\text{WN}}(T)$)

La famille de termes $(\text{RED}^{\text{WN}}(T))_{T \in \text{Type}}$ est définie par induction sur les types :

- $\text{RED}^{\text{WN}}(X) = \text{Norm}(X)$;
- $\text{RED}^{\text{WN}}(U \rightarrow V) = \{t : U \rightarrow V; \forall u \in \text{RED}^{\text{WN}}(U), (t)u \in \text{RED}^{\text{WN}}(V)\}.$

2.2 Énoncé du théorème et structure de la preuve

La normalisation (faible) est conséquence des deux lemmes suivants :

Lemme 2.4 (*Adaptation*)

Pour tout type T , on a $\text{Neut}(T) \subseteq \text{RED}^{\text{WN}}(T) \subseteq \text{Norm}(T)$.

Lemme 2.5 (*Adéquation*)

Soit $t : U$ de variables libres parmi $x_1^{T_1}, \dots, x_n^{T_n}$.
Pour tous $(u_i \in \text{RED}^{\text{WN}}(T_i))_{1 \leq i \leq n}$, on a $t \{u_i/x_i\} \in \text{RED}^{\text{WN}}(U)$.

À partir de ces deux lemmes, le résultat est très facile :

Théorème 2.6

Tout terme simplement typé est (faiblement) normalisable.

Démonstration : Soit $t : T$ de variables libres $(x_i^{T_i})_{1 \leq i \leq n}$. Le lemme d'adaptation (2.4) nous assure que pour tout $1 \leq i \leq n$, $x_i^{T_i} \in \text{RED}^{\text{WN}}(T_i)$ puisque les variables sont neutres.

Le lemme d'adéquation nous assure alors que $t \{x_i^{T_i}/x_i, 1 \leq i \leq n\} = t$ est réductible de type T .

Or, en utilisant à nouveau le lemme d'adaptation, on a $t \in \text{RED}^{\text{WN}}(T) \subseteq \text{Norm}(T)$ ce qui permet de conclure que t est normalisable. □

2.3 Preuve des deux lemmes clés

Le lemme d'adaptation repose sur le fait suivant :

Lemme 2.7

Pour tout terme t et toute variable x , si $(t)x$ est normalisable, alors t l'est également.

Remarque 2.8

On a énoncé le lemme précédent sans indication de type. En effet, le résultat est vrai dans le λ -calcul pur.

Démonstration : Soit $(t)x \rightarrow^* n$ où n est en forme normale. De deux choses l'une : soit la réduction se passe intégralement dans t , et dans ce cas $n = (n')x$ avec $t \rightarrow^* n'$ d'où t est normalisable, soit $(t)x \rightarrow^* (\lambda x^T.u)x \rightarrow^* u \rightarrow^* n$ et dans ce cas, $t \rightarrow^* \lambda x^T.u \rightarrow^* \lambda x^T.n$ et t est également normalisable. □

On a maintenant les outils pour démontrer le lemme d'adaptation :

Démonstration du lemme 2.4 : La démonstration se fait par induction sur la structure du type T :

- Si $T = X$, alors $\text{RED}^{\text{WN}}(X) = \text{Norm}(X) \supseteq \text{Neut}(X)$ par définition.
- Si $T = U \rightarrow V$, on a alors :
 - $\text{Neut}(T) \subseteq \text{RED}^{\text{WN}}(T)$: soit $t \in \text{Neut}(T)$, c'est-à-dire qu'il soit neutre et normal de type $U \rightarrow V$. Soit $u \in \text{RED}^{\text{WN}}(U)$, $(t)u$ est neutre. Par ailleurs, u étant normalisable, soit v sa forme normale, on a $(t)v$ neutre et normal donc $(t)v \in \text{neutre}V \subseteq \text{RED}^{\text{WN}}(V)$ (par hypothèse d'induction) et $(t)u \in \text{RED}^{\text{WN}}(V)$ par clôture des réductibles par β -expansion.
 - $\text{RED}^{\text{WN}}(T) \subseteq \text{Norm}(T)$: soit $t \in \text{RED}^{\text{WN}}(T)$, montrons qu'il est normalisable. On a $x^U \in \text{Neut}(U) \subseteq \text{RED}^{\text{WN}}(U)$ (par hypothèse d'induction) et le lemme précédent nous permet de conclure que $t \in \text{Norm}(T)$. □

Le lemme d'adéquation repose sur le lemme suivant, qui énonce qu'on ne peut pas “devenir réductible” par réduction : si une séquence de β -réduction aboutit dans $\text{RED}^{\text{WN}}(T)$, c'est que le terme dont elle était issue en était déjà membre. (la β -expansion et la relation inverse de la β -réduction) :

Lemme 2.9

Pour tout type T , $\text{RED}^{\text{WN}}(T)$ est clos par anti-réduction (ou β -expansion) :

Si $t \in \text{RED}^{\text{WN}}(T)$ et $u \rightarrow_{\beta} t$, alors $u \in \text{RED}^{\text{WN}}(T)$.

Démonstration : On prouve le lemme par induction sur le type T .

- Le cas de base est trivial puisque si T est un type atomique, $\text{RED}^{\text{WN}}(T) = \text{Norm}(T)$ et que la normalisabilité est une propriété close par anti-réduction.
- Si $T = U \rightarrow V$, alors soit $t \in \text{RED}^{\text{WN}}(T)$ et $t' \rightarrow t$: on veut montrer que $t' \in \text{RED}^{\text{WN}}(T)$. Soit alors $u \in \text{RED}^{\text{WN}}(U)$. On a $(t)u \in \text{RED}^{\text{WN}}(V)$ par définition et l'hypothèse d'induction sur V nous assure que $\text{RED}^{\text{WN}}(V)$ est clos par anti-réduction donc que $(t')u \in \text{RED}^{\text{WN}}(V)$ et ceci pour tout $u \in \text{RED}^{\text{WN}}(U)$: on en conclut que $t' \in \text{RED}^{\text{WN}}(T)$ ce qui achève la preuve du lemme. □

On peut alors démontrer le lemme d'adéquation :

Démonstration du lemme 2.5 : On raisonne par induction sur la structure de $t : T$.

- Si $t = x_i^{T_i}$, le résultat est trivial car u_i est dans $\text{RED}^{\text{WN}}(T_i)$ par hypothèse du lemme.
- Si $t = \lambda x^U . t'$ avec $T = U \rightarrow V$. Soit $u \in \text{RED}^{\text{WN}}(U)$, on considère $w = (\lambda x^T . t')u$. Par hypothèse d'induction, et en supposant les x_1 non libres dans $u, u_1, \dots, u_n$, on a $t'\{u_1/x_1, \dots, u_n/x_n, u/x\} \in \text{RED}^{\text{WN}}(V)$ donc par le lemme précédent, $(\lambda x^T . t'\{u_i/x_i, 1 \leq i \leq n\})u = w\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(V)$ ce qui montre que $t\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(T)$.
- Si $t = (u)v$, avec $u : V \rightarrow T$ et $v : V$. Par hypothèse d'induction, $u' = u\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(V \rightarrow T)$ et $v' = v\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(V)$ d'où $t\{u_i/x_i, 1 \leq i \leq n\} = (u')v' \in \text{RED}^{\text{WN}}(T)$ ce qui conclut la preuve du lemme d'adéquation. □

Ceci achève la preuve de normalisation faible du λ -calcul simplement typé.

2.4 Vers la normalisation forte

On peut adapter la notion de réductibilité pour obtenir une preuve de la normalisation forte.

Il faut cependant adapter la réductibilité à la propriété de normalisation forte. En effet :

- Contrairement à WN, SN n'est pas stable par β -expansion en général (sinon SN et WN seraient équivalents simplement parce qu'une forme normale est toujours SN et tout terme normalisable est le β -expansé d'une forme normale).
- En revanche, SN est stable par β -réduction (contrairement à WN...)
- On avait auparavant la propriété $\text{Neut}(T) \subseteq R \subseteq \text{Norm}(T)$. Il va nous falloir raffiner cet encadrement de manière à être plus précis, par exemple en demandant que $R \subseteq \text{SN}(T)$.

3 Théorème de normalisation forte de $\Lambda^{\rightarrow}$

Pour prouver la normalisation forte, on va suivre le même schéma que dans la section précédente en raffinant la notion de réductibilité pour prouver que tout terme réductible est en fait fortement normalisable. Un résultat d'adéquation comme ci-dessus permettra alors de conclure.

3.1 Neutralité, réductibilité

Pour la normalisation faible, on utilisait un encadrement des ensembles de termes réductibles par deux ensembles : les termes neutres et les normalisables de terme T :

$$\text{Neut}(T) = \{t \in \Lambda^{\rightarrow}; t \text{ normal et neutre de type } T\}$$

$$\text{Norm}(T) = \{t \in \Lambda^{\rightarrow}; t \text{ normalisable de type } T\}.$$

Pour la preuve de normalisation forte, on va raffiner l'encadrement du lemme d'adaptation en considérant deux variantes des ensembles de termes ci-dessus : on remplace l'ensemble des termes neutres et normaux par l'ensemble des termes neutres se réduisant uniquement vers des termes réductibles et on remplace l'ensemble $\text{Norm}(T)$, des termes normalisables de type T par l'ensemble $\text{SNorm}(T)$ des termes fortement normalisables de type T .

Si la définition du premier ensemble est repoussé après la définition de la réductibilité, puisqu'elle en dépend, on définit ci-dessous $\text{SNorm}(T)$:

Définition 3.1 ($\text{SNorm}(T)$)

$$\text{SNorm}(T) = \{t \in \Lambda^\rightarrow; t \text{ fortement normalisable de type } T\}.$$

Par ailleurs, à chaque type T on associe un nouvel ensemble de termes réductibles de type T :

Définition 3.2 ($\text{RED}^{\text{SN}}(T)$)

La famille de termes $(\text{RED}^{\text{SN}}(T))_{T \in \text{Type}}$ est définie par induction sur les types :

- $\text{RED}^{\text{SN}}(X) = \text{SNorm}(X)$
- $\text{RED}^{\text{SN}}(U \rightarrow V) = \{t : U \rightarrow V; \forall u \in \text{RED}^{\text{SN}}(U), (t)u \in \text{RED}^{\text{SN}}(V)\}.$

Remarque 3.3

On notera que par rapport à la réductibilité pour la normalisation faible, on a simplement modifié la clause des types atomiques, mais que bien évidemment cela impacte l'ensemble de la définition. On notera par ailleurs qu'une fois qu'on a obtenu le théorème de normalisation, on saura que les deux notions coïncident en fait... mais patientons !

On peut maintenant définir notre notion de neutralité adaptée à la normalisation forte :

Définition 3.4 ($\text{Neut}^{\text{SN}}(T)$)

$$\text{Neut}^{\text{SN}}(T) = \{t \in \Lambda^\rightarrow; t \text{ est neutre de type } T \text{ et } \forall t', t \rightarrow_\beta t', t' \in \text{RED}^{\text{SN}}(T)\}$$

On peut alors noter, de manière immédiate :

Proposition 3.5

Pour tout type T , on a :

- $\text{Neut}(T) \subseteq \text{Neut}^{\text{SN}}(T)$;
- $\text{SNorm}(T) \subseteq \text{Norm}(T)$.

Démonstration : Si $t \in \text{Neut}(T)$, alors t étant normal, l'ensemble de ses réduits est vide et satisfait donc la propriété demandée.

Si $t \in \text{SNorm}(T)$, alors puisqu'il est fortement normalisable il est normalisable et $t \in \text{Norm}(T)$. □

3.2 Énoncé du théorème et structure de la preuve

La normalisation forte est, presque comme précédemment, conséquence des deux lemmes suivants :

Lemme 3.6 (*Adaptation*)

Pour tout type T , on a $\text{Neut}^{\text{SN}}(T) \subseteq \text{RED}^{\text{SN}}(T) \subseteq \text{SNorm}(T)$.

Lemme 3.7 (*Adéquation*)

Soit $t : U$ de variables libres parmi $x_1^{T_1}, \dots, x_n^{T_n}$. Pour tous $(u_i \in \text{RED}^{\text{SN}}(T_i))_{1 \leq i \leq n}$, on a $t\{u_i/x_i\} \in \text{RED}^{\text{SN}}(U)$.

À partir de ces deux lemmes, le résultat est très facile :

Théorème 3.8

Tout terme simplement typé est fortement normalisable.

Démonstration : Soit $t : T$ de variables libres $(x_i^{T_i})_{1 \leq i \leq n}$. Le lemme d'adaptation (3.6) nous assure que pour tout $1 \leq i \leq n$, $x_i^{T_i} \in \text{RED}^{\text{SN}}(T_i)$ puisque les variables de type T sont neutres et sont donc dans $\text{Neut}^{\text{SN}}(T)$.

Le lemme d'adéquation (3.7) nous assure alors que $t \{x_i^{T_i}/x_i, 1 \leq i \leq n\} = t$ est réductible de type T ($\text{RED}^{\text{SN}}(T)$).

Or, en utilisant à nouveau le lemme d'adaptation, on a $t \in \text{RED}^{\text{SN}}(T) \subseteq \text{SNorm}(T)$ ce qui permet de conclure que t est fortement normalisable. $\square$

3.3 Preuve des deux lemmes clés

Le lemme d'adaptation repose sur le fait suivant :

Lemme 3.9

Pour tout type T , $\text{RED}^{\text{SN}}(T)$ est clos par β -réduction :

$$t \in \text{RED}^{\text{SN}}(T), \quad t \longrightarrow_{\beta} t' \quad \Rightarrow \quad t' \in \text{RED}^{\text{SN}}(T).$$

Démonstration : On prouve le lemme par induction sur le type T .

- Le cas de base est trivial puisque si T est un type atomique, $\text{RED}^{\text{SN}}(T) = \text{SNorm}(T)$ et que la forte normalisation est une propriété close par β -réduction.
- Si $T = U \rightarrow V$, alors soit $t \in \text{RED}^{\text{SN}}(T)$ et $t' \longrightarrow t$: on veut montrer que $t' \in \text{RED}^{\text{SN}}(T)$. Soit alors $u \in \text{RED}^{\text{SN}}(U)$. On a $(t)u \in \text{RED}^{\text{SN}}(V)$ par définition et l'hypothèse d'induction sur V nous assure que $\text{RED}^{\text{SN}}(V)$ est clos par β -réduction donc que $(t')u \in \text{RED}^{\text{SN}}(V)$ et ceci pour tout $u \in \text{RED}^{\text{SN}}(U)$: on en conclut que $t' \in \text{RED}^{\text{SN}}(T)$ ce qui achève la preuve du lemme. $\square$

Lemme 3.10

Pour tout terme t et toute variable x , si $(t)x$ est fortement normalisable, alors t l'est également.

Remarque 3.11

On a énoncé le lemme précédent sans indication de type. En effet, le résultat est vrai dans le λ -calcul pur.

Ce résultat est en fait trivial et présenté ici surtout pour renforcer le parallèle avec la preuve de normalisation faible de la section précédente :

Démonstration : Si t n'était pas fortement normalisant, on aurait une réduction infinie issue de t , $(t_i)_{i \in \mathbb{N}}$, d'où on obtiendrait immédiatement une réduction infinie issue de $(t)x$, $((t_i)x)_{i \in \mathbb{N}}$. Par contraposé, la normalisation forte de $(t)x$ entraîne celle de t . $\square$

On a maintenant les outils pour démontrer le lemme d'adaptation :

Démonstration du lemme 3.6 : La démonstration se fait par induction sur la structure du type T :

- Si $T = X$, alors $\text{RED}^{\text{SN}}(X) = \text{SNorm}(X)$ par définition. Par ailleurs, si $t \in \text{Neut}^{\text{SN}}(X)$, alors pour tout t' tel que $t \longrightarrow_{\beta} t'$, on a $t' \in \text{RED}^{\text{SN}}(X) = \text{SNorm}(X)$ donc t est fortement normalisable et $t \in \text{RED}^{\text{SN}}(X)$.
- Si $T = U \rightarrow V$, on a alors :
 - $\text{Neut}^{\text{SN}}(T) \subseteq \text{RED}^{\text{SN}}(T)$: soit $t \in \text{Neut}^{\text{SN}}(T)$, c'est-à-dire qu'il soit neutre et que tous ses réduits sont réductibles de type $U \rightarrow V$: $\forall t', t \longrightarrow_{\beta} t'$, on a $t' \in \text{RED}^{\text{SN}}(T)$.
On montre par induction sur la longueur de la plus longue réduction issue de u , $\ell(u)$, que pour tout $u \in \text{RED}^{\text{SN}}(U) \subseteq \text{SNorm}(U)$, $(t)u \in \text{RED}^{\text{SN}}(V)$.
 - si $\ell(u) = 0$, alors u est normal, et les réduits de $(t)u$ sont donc de la forme $(t')u$ où t' est un réduct de t donc $t' \in \text{RED}^{\text{SN}}(U \rightarrow V)$ et $(t')u \in \text{RED}^{\text{SN}}(V)$. En conséquence, tous les réduits de $(t)u$, qui est neutre, sont dans $\text{RED}^{\text{SN}}(V)$: $(t)u \in \text{Neut}^{\text{SN}}(V) \subseteq \text{RED}^{\text{SN}}(V)$.

- si la propriété est vraie des termes v pour lesquels $\ell(v) \leq n$ et supposons $\ell(u) = n + 1$. Soit v un réduit de $(t)u$, il est soit obtenu en réduisant t en un $t' : v = (t')u$, soit obtenu en réduisant u en u' , $v = (t)u'$, mais il n'y a pas d'autre possibilité puisque t est neutre. Dans le premier cas, comme on sait que $t' \in \text{RED}^{\text{SN}}(U \rightarrow V)$, on a $v \in \text{RED}^{\text{SN}}(V)$. Dans le second cas, comme $\ell(u') \leq n$, et $u' \in \text{RED}^{\text{SN}}(U)$ par clôture de $\text{RED}^{\text{SN}}(U)$ par réduction, l'hypothèse d'induction nous assure que $v \in \text{RED}^{\text{SN}}(V)$. Tous les réduits de $(t)u$ étant dans $\text{RED}^{\text{SN}}(V)$ et $(t)u$ étant neutre, on en déduit que $(t)u \in \text{Neut}^{\text{SN}}(V) \subseteq \text{RED}^{\text{SN}}(V)$.
- $\text{RED}^{\text{SN}}(T) \subseteq \text{SNorm}(T)$: soit $t \in \text{RED}^{\text{SN}}(T)$, montrons qu'il est fortement normalisable. On a $x^U \in \text{Neut}(U) \subseteq \text{Neut}^{\text{SN}}(U) \subseteq \text{RED}^{\text{SN}}(U)$ (par hypothèse d'induction) donc $(t)x \in \text{RED}^{\text{SN}}(V) \subseteq \text{SNorm}(V)$ et le lemme précédent nous permet de conclure que $t \in \text{Norm}(T)$.

□

Le lemme d'adéquation repose sur :

Lemme 3.12

$$(\forall u \in \text{RED}^{\text{SN}}(U), v\{u/x\} \in \text{RED}^{\text{SN}}(V)) \Rightarrow \forall u \in \text{RED}^{\text{SN}}(U), (\lambda x.v)u \in \text{RED}^{\text{SN}}(V).$$

et son corollaire immédiat par définition de $\text{RED}^{\text{SN}}(U \rightarrow V)$:

Corollaire 3.13

$$(\forall u \in \text{RED}^{\text{SN}}(U), v\{u/x\} \in \text{RED}^{\text{SN}}(V)) \Rightarrow \lambda x.v \in \text{RED}^{\text{SN}}(U \rightarrow V).$$

Démonstration du lemme : Soit $v : V$ ayant possiblement x^U comme variable libre, et supposons que pour tout $u \in \text{RED}^{\text{SN}}(U)$, $v\{u/x\} \in \text{RED}^{\text{SN}}(V)$.

On remarque tout d'abord que, les variables étant réductibles pour leur type, $v = v\{x^U/x\} \in \text{RED}^{\text{SN}}(V)$. En particulier, u et v sont fortement normalisants. On montre alors le résultat par induction sur $\ell(u) + \ell(v)$, en considérant $t = (\lambda x.v)u$ qui est neutre, il suffit de montrer qu'il est dans $\text{Neut}^{\text{SN}}(V)$ pour avoir le résultat par le lemme d'adaptation, on va donc considérer ses réduits et montrer qu'ils sont tous réductibles de type V . Soit t' un réduit de t . Alors de trois choses l'une :

- $t' = v\{u/x\}$ qui est dans $\text{RED}^{\text{SN}}(V)$ par hypothèse du lemme ;
- $t' = (\lambda x.v')u$ avec v' un réduit de v . On a alors $v' \in \text{RED}^{\text{SN}}(V)$ et pour tout $u \in \text{RED}^{\text{SN}}(U)$, $v'\{u/x\} \in \text{RED}^{\text{SN}}(V)$ puisque $v\{u/x\} \rightarrow_{\beta} v'\{u/x\}$ et que $\text{RED}^{\text{SN}}(V)$ est clos par réduction. Comme $\ell(u) + \ell(v') < \ell(u) + \ell(v)$, on peut appliquer l'hypothèse d'induction et conclure que $t' = (\lambda x.v')u \in \text{RED}^{\text{SN}}(V)$.
- $t' = (\lambda x.v)u'$ avec u' un réduit de u qui est donc dans $\text{RED}^{\text{SN}}(U)$, on sait alors par hypothèse que $v\{u'/x\} \in \text{RED}^{\text{SN}}(V)$. Puisque $\ell(u') + \ell(v) < \ell(u) + \ell(v)$, on peut appliquer l'hypothèse d'induction et conclure que $t' \in \text{RED}^{\text{SN}}(V)$.

□

On peut alors démontrer le lemme d'adéquation :

Démonstration du lemme 2.5 : On raisonne par induction sur la structure de $t : T$.

- Si $t = x_i^{T_i}$, le résultat est trivial car u_i est dans $\text{RED}^{\text{SN}}(T_i)$ par hypothèse du lemme.
- Si $t = \lambda x^U.t'$ avec $T = U \rightarrow V$ et en choisissant x^U pour qu'elle ne soit libre dans aucun des u_i et qu'elle soit distincte de tous les $x_i^{T_i}$. Soit $u \in \text{RED}^{\text{SN}}(U)$. Par hypothèse d'induction, $t'\{u_i/x_i, 1 \leq i \leq n\}\{u/x\} = t'\{u_i/x_i, 1 \leq i \leq n, u/x\} \in \text{RED}^{\text{SN}}(V)$ donc par le lemme précédent, on a $t\{u_i/x_i, 1 \leq i \leq n\} = \lambda x.t'\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{SN}}(U \rightarrow V)$.
- Si $t = (u)v$, avec $u : V \rightarrow T$ et $v : V$. Par hypothèse d'induction, $u' = u\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(V \rightarrow T)$ et $v' = v\{u_i/x_i, 1 \leq i \leq n\} \in \text{RED}^{\text{WN}}(V)$ d'où $t\{u_i/x_i, 1 \leq i \leq n\} = (u')v' \in \text{RED}^{\text{WN}}(T)$ ce qui conclut la preuve du lemme d'adéquation.

□

Ceci achève la preuve de normalisation forte du λ -calcul simplement typé.