Cours de Lambda-calcul[*]

Ch. Berline

CHAPITRE 2 : Syntaxe du λ-calcul non typé.
(2. fnt, termes résolubles, stratégies de réduction, normalisation et standardisation).[†]

Table des matières

1 Réduction de tête et réduction gauche.

L'ensemble N des *termes normaux* est le plus petit ensemble qui contient les variables et qui est tel que $\lambda\bar{x}.y\bar{t}$ est dans N dès que les $t_i$sont dans N ($\bar{x}$,$y \in V$, $l(\bar{x}), l(\bar{t}) \geq 0$). Une classe plus générale est celle des *"formes normales de tête"* ou *"fnt"*.

Dfinition 1 *Une* forme normale de tête *est un terme de la forme $\lambda\bar{x}.y\bar{t}$; on parlera aussi de terme* normal de tête.

La différence avec les termes normaux étant qu'il n'y a aucune condition sur les t_i. En fait un terme t est normal si et seulement si il est "héréditairement normal de tête".

Il est très facile de voir, par induction sur la longueur des termes, que tout terme t est soit normal de tête soit de la forme $\lambda\bar{x}.R\bar{t}$, où R est un β-redex ; dans ce dernier cas R est appelé *le redex de tête de t*. Dans la suite de ce chapitre, et sauf précision contraire, *redex* sera synonyme de β-redex.

[*]Cours du DEA de Logique et Fondements de l'Informatique de l'Université Paris 7.
[†]Version du 5 Février 2002+

Remarque 2 *1. Le redex de tête, si il existe, est manifestement le redex le plus à gauche de t, c'est à dire le redex dont le "$(\lambda$" est le plus à gauche de t; la réciproque est fausse : par exemple $(x)R$, où R est un redex, n'a pas de redex de tête, et R est évidemment son redex le plus à gauche.*
2. Tout β ou $\beta\eta$ réduit d'un terme normal de tête est lui-même normal de tête.

Une étape de réduction de tête consiste à réduire le rédex de tête (si il y en a un). La notation $t \to_h t'$ signifiera dans ce cours[1] qu'on peut atteindre t' à partir de t par une succession de réductions de tête.

Dfinition 3 La réduction gauche *(sous-entendu : maximale) est la stratégie qui consiste à réduire toujours le redex le plus à gauche d'un terme non normal, et qui s'arrète quand elle rencontre un terme normal.* La réduction de tête *(maximale) est la stratégie qui réduit toujours le redex de tête, si il y en a un, et s'arrête sinon.*

Dfinition 4 *Un terme est* normalisable à gauche *ou* g-normalisable *si sa réduction gauche termine. Un terme est* normalisable de tête *ou* h-normalisable *si sa réduction de tête termine.*

Puisque le redex de tête est, quand il existe, le redex le plus à gauche, la réduction gauche d'un terme commence nécessairement par sa réduction de tête, et coïncide avec elle si celle-ci est infinie ; en particulier tout terme normalisable à gauche est nécessairement normalisable de tête.

Exemple 5 $\omega \equiv \delta\delta$ *, with $\delta \equiv \lambda x.xx$ est l'exemple le plus classique d'un terme qui n'est pas normalisable de tête (et qui n'est donc pas normalisable à gauche), en effet :*

$$\omega \equiv \delta\delta \to_h \delta\delta \to_h \delta\delta \to_h \ldots\ldots$$

$F\omega$ est h- et g- normalisable (sa réduction de tête se termine sur I), mais il n'est pas fortement normalisable.

$x\omega$ est normal de tête mais pas normalisable.

YK et $K^\infty \equiv (\lambda x.K(x)x)\lambda x.K(x)x$ sont d'autres exemples intéressants de termes clos non normalisables de tête : cette fois la réduction de tête produit une infinité de "λ de tête", puisque :

$$YK \to_h K^\infty \to_h \lambda x_1.K^\infty \to_h \lambda x_1\lambda x_2.K^\infty \to_h \ldots$$

Remarque 6 *Il est évident que t est h-normalisable (resp. g-normalisable) ssi $\lambda\bar{x}.t$ l'est, ceci indépendamment du choix de $\bar{x}$.*

L'importance de la réduction gauche tient dans ses bonnes propriétés mathématiques ; en particulier, et surtout, dans le théorème suivant qui montre que la stratégie gauche est une stratégie sûre pour découvrir la forme normale d'un terme normalisable.

[1] Note : dans Barendregt cette notation a un sens différent (une étape de réduction de tête !).

Thorme 7 ***de normalisation gauche.***
normalisable$\Longleftrightarrow$*normalisable à gauche*

Preuve. L'implication non triviale sera prouvée plus tard.

L'inconvénient de la stratégie gauche c'est qu'elle multiplie le calcul des arguments : par exemple la réduction gauche de δv, avec $\delta = \lambda x.xx$ et $v \equiv Ix$, calculera v deux fois.

**Appel par nom, appel par valeur.*
La stratégie de réduction gauche est aussi appelée "*stratégie de l'appel par nom*", en contraste avec la "*stratégie d'appel par valeur*" qui calcule toujours un argument avant de le donner à une fonction. Dans cette dernière stratégie seuls les redex de la forme $(\lambda x.t)u$, où u appartient à un certain ensemble de termes considérés comme des "valeurs" peuvent être réduits. L'appel par valeur peut être simulé dans l'appel par nom.

**Réduction paresseuse.* Toute stratégie a une version dite *paresseuse*, qui s'arrête dès que l'on atteint une abstraction (l'idée étant que c'est inutile de calculer une fonction qui ne sera jamais appliquée à un argument). Les langages concrets de programmation fonctionnelle sont en général paresseux. La définition suivante est particulièrement utile dans ce cadre. Remarquer que la réduction gauche paresseuse d'un terme clos coïncide toujours avec la réduction de tête paresseuse, puisque dans les deux cas on s'arrête à la première abstraction.

Dfinition 8 L'ordre de fonctionalité $o(t)$ *d'un terme t est le plus grand entier $n \geq 0$ tel que t peut se β-réduire à un terme de la forme $\lambda\bar{x}.u$ avec $l(\bar{x}) = n$, si il n'existe pas de tel entier on pose $o(t) \equiv \infty$.*

Par exemple $o(x) = 0$, $o(II) = 1$, $o(K) = 2$, $o(K^\infty) = \infty$.
Les termes clos d'ordre zéro sont non normalisable de tête ; la raison en est qu'ils sont nécessairement de la forme $(u)v$, avec u et v clos, et tous leurs réduits sont de la même forme, et en particulier ne sont pas des fnt.

Commentaires sur les réductions gauches et de tête.
Supposons que la réduction de tête de t se termine sur $\lambda\bar{x}.y\bar{t}$. Alors, ou le terme est normal et on s'arrête, ou la réduction gauche de t continue dans les sous-termes t_i (de gauche à droite) et, quoiqu'il arrive par la suite, tous les réduits successifs seront de la forme $\lambda\bar{x}.y\bar{u}$ avec $l(\bar{t}) = l(\bar{u})$ $(= n)$. A la fin de la réduction de tête nous avons ainsi obtenu un invariant, à savoir $(\lambda\bar{x}.y\,,\,n)$, qui peut être vu comme un résultat partiel du calcul. Au cours de la suite éventuelle de la réduction gauche de t nous récupèrerons peut-être d'autres informations semblables, que nous pourrons rassembler.

Si la réduction gauche termine nous aurons un résultat complet : la forme normale du terme.

Si la réduction gauche ne termine pas, il peut se produire, sans que cela soit obligé, que l'on continue à recevoir indéfiniment des informations partielles (analogie : un programme qui sortirait la liste des nombres premiers).

Si t n'est pas normalisable de tête, sa réduction de tête, qui coincide dans ce cas avec sa réduction gauche, durera indéfiniment, sans jamais produire aucune information sur t, si ce n'est éventuellement une suite croissante (finie ou non) de minorants de son ordre de fonctionnalité. C'est la raison pour laquelle les termes *non normalisables de tête* sont considérés comme représentant les programmes qui tournent indéfiniment sans produire d'information (par exemple parce qu'ils bouclent [2]). Pour la même raison ils sont aussi considérés, dans d'autres contextes, comme représentant *l'indéfini*, ou le *manque d'information* (pour les versions paresseuses on est plus drastique et on ne prend que ceux d'ordre zéro).

On pourrait penser à prendre la non normalisabilité pour représenter l'indéfini, mais, outre les remarques "opérationnelles" ci-dessus qui votent en faveur de la non normalisabilité de tête, il y a aussi des raisons mathématiques de ne pas le faire. En effet il est possible d'égaliser de manière consistante tous les termes non h-normalisables, comme nous le verrons dans la suite du cours, mais ceci est faux pour les non normalisables, comme le montre l'exercice suivant.

Exercice 9 *Une* λ-congruence *est une relation d'équivalence sur* Λ, *qui est contextuelle et qui contient* $=_\beta$. *Montrer qu'il n'existe pas de* λ*-congruence non triviale qui rende tous les termes non-normalisables équivalents.*[3]

Exercice 10 *Si* t *et* t' *sont* β*-équivalents et h-normalisables, alors le couple* $(\lambda\bar{x}.y\,,\,n)$ *qui est donné par leur réduction de tête est le même dans les deux cas.*

Lemme 11 *1.* t *est normalisable de tête ssi* $\lambda x.t$ *l'est*
2. t *est normalisable de tête ssi* tx *l'est.*

Preuve. 1. suit immédiatement de la définition de h-normalisable.

2. Si $t = \lambda x.u$ alors $tx \to_h u$ donc tx est normalisable de tête ssi u l'est ssi t l'est. Si t n'est pas une abstraction la réduction de tête de tx commence dans t. Supposons que t ait un réduit de tête qui est une abstraction, soit $\lambda x.u$, qui apparaisse à la nième étape ($n \geqslant 1$), alors la réduction de tête de tx commence par $tx \to_h^n (\lambda x.u)x \to_h u$ et se continue par celle de u. Comme précédemment on en déduit que tx est normalisable de tête ssi u l'est, ssi $\lambda x.u$ l'est, ssi t l'est. Par ailleurs si t n'est pas une abstraction et si sa réduction de tête n'en fait pas apparaître, alors la réduction de tête de tx se passe entièrement dans t et donc tx est normalisable de tête ssi t l'est.

Lemme 12 *La réduction de tête est substitutive. Autrement dit :*
$u \to_h u'$ *implique* $u[\bar{x} := \bar{v}] \to_h u'[\bar{x} := \bar{v}]$.

---

[2](en anglais : *looping)*

[3]Suggestion : regarder les termes $\lambda x.x\omega V$ et $\lambda x.x\omega F$.

Preuve. C'est une conséquence immédiate du lemme suivant.

Notation 13 *On désigne par w^k le terme obtenu après k étapes de réduction de tête de w, si la réduction de tête de w admet au moins k étapes.*

Lemme 14 *Si u n'est pas normal de tête alors $u[\bar{x} := \bar{v}]$ n'est pas normal de tête, et pour tout v on a $u[\bar{x} := \bar{v}]^1 = u^1[\bar{x} := \bar{v}]$.*

Preuve. Traitons d'abord le cas où $u = (\lambda y.t)r$ est un redex. On suppose $y \notin \bar{x} \cup Vlib(\bar{v})$. Alors $u[\bar{x} := \bar{v}] =_{def} (\lambda y.t[\bar{x} := \bar{v}])r[\bar{x} := \bar{v}]$. On a donc
$u[\bar{x} := \bar{v}]^1 =_{def} t[\bar{x} := \bar{v}][y := r[\bar{x} := \bar{v}]] =$
$= t[\bar{x} := \bar{v},\ y := r[\bar{x} := \bar{v}]]$ (car $y \notin \bar{x} \cup Vlib(\bar{v})$) $=$
$= t[y := r[\bar{x} := \bar{v}], \bar{x} := \bar{v}] =$
$= t[y := r][\bar{x} := \bar{v}] =$
$= u^1[\bar{x} := \bar{v}]$ q.e.d.
Supposons maintenant que $u = \lambda\bar{y}.R\bar{w}$, pour un certain redex R. Quitte à renommer les $\bar{y}$ on suppose $\bar{y} \cap \bar{x} = \emptyset$. Alors $u[\bar{x} := \bar{v}] = \lambda\bar{y}.\,R[\bar{x} := \bar{v}]\bar{w}[\bar{x} := \bar{v}]$. On a :
$u[\bar{x} := \bar{v}]^1 =_{def} \lambda\bar{y}.\,R[\bar{x} := \bar{v}]^1\bar{w}[\bar{x} := \bar{v}] =$
$= \lambda\bar{y}.\,R^1[\bar{x} := \bar{v}]\bar{w}[\bar{x} := \bar{v}]$ (cas précédent)=
$= \lambda\bar{y}.\,(R^1\bar{w})[\bar{x} := \bar{v}]$ (def de la substitution sur les applications)
$= (\lambda\bar{y}.\,R^1\bar{w})[\bar{x} := \bar{v}]$ (def de la substitution sur les abstractions)
$= u^1[\bar{x} := \bar{v}]$.

Corollaire 15 *Si $u[\bar{x} := \bar{v}]$ est h-normalisable, alors u est h-normalisable.*

Preuve. On déduit du lemme précédent que si u a au moins n réduits de tête, alors $u[\bar{x} := \bar{v}]$ en a au moins n également et $u[\bar{x} := \bar{v}]^n = u^n[\bar{x} := \bar{v}]$ (induction sur n). L'assertion suit.

2 Termes résolubles.

Il existe une *caractérisation algébrique* des termes normalisables de tête, en tant que *termes résolubles*, en un sens que nous allons définir bientôt. Le fait que les deux notions soient équivalentes n'est pas du tout trivial, en fait il s'agit d'un théorème essentiellement équivalent au théorème de normalisation gauche, en ce sens qu'on peut démontrer les deux avec des arguments voisins, ou encore déduire assez facilement l'un de l'autre, dans les deux sens. Une fois que l'équivalence des deux notions est prouvée on a le choix d'utiliser librement la vue opérationnelle ou la vue algébrique suivant le problème à résoudre.

Dfinition 16 *On appelle clôture d'un terme t tout terme $\lambda\bar{x}.t$ tel que $\bar{x} = Vlib(t)$.*

Les clôtures de t ne diffèrent donc que par l'ordre dans lequel $Vlib(t)$ est énuméré sous la forme $\bar{x}$, un terme clos a pour seule clôture lui-même.

Dfinition 17 *Un terme t est* résoluble *si il existe une clôture t^c de t et une suite $\bar{v}$ tels que $t^c\bar{v} =_\beta I$.*

En particulier un terme *clos* t est *résoluble* si et seulement si il existe $\bar{v}$ tel que $t\bar{v} =_\beta I$.

Remarque 18 *Il résulte immédiatement de la définition que :*

1. *On peut remplacer $=_\beta$ par $\rightarrow_\beta$.*
2. *On peut remplacer I par x.*
3. *On peut exiger $l(\bar{v}) \geqslant n$, pour tout n fixé arbitrairement.*
4. *On peut exiger : "pour toute clôture t^c de t il existe ...".*
5. *Si t est résoluble, alors pour tout terme w il existe $\bar{v}$ tel que $t^c\bar{v} =_\beta w$.*

Pour le point 3 ci-dessus il suffit de remarquer que si $u =_\beta I$ alors $uII...II =_\beta I$. Pour le point 4 on peut argumenter de la façon suivante. Soient $t^c, \bar{v}$ tels que t^c est une clôture de t et $t^c\bar{v} =_\beta I$. Par le point 3 on peut supposer $l(\bar{v}) \geq l(\bar{x})$ où $\bar{x} = Vlib(t)$. Pour toute autre clôture $t^{c\prime}$ de t on a alors $t^{c\prime}\bar{v}' =_\beta t^c\bar{v}$ pour $\bar{v}'$ une permutation adéquate des éléments de $\bar{v}$, d'où l'affirmation.

Proposition 19 *t est résoluble ssi il existe $\bar{y}, \bar{u}, \bar{v}$ tels que $t[\bar{y} := \bar{u}]\bar{v} =_\beta I$*

Preuve. $\Rightarrow$ Si t est résoluble il existe $\bar{v}$ tel que $t^c\bar{v} =_\beta I$ et l'on peut supposer que $l(\bar{v}) \geqslant l(\bar{y})$ où $\bar{y}$ est tel que $t^c = \lambda\bar{y}.t$ (autrement dit $\bar{y}$ est la suite des variables libres de t prises dans un certain ordre).Soient $\bar{u}$ et $\bar{w}$ tels que $\bar{v} = \bar{u}\bar{w}$ et $l(\bar{u}) = l(\bar{y})$; on a alors $t[\bar{y} := \bar{u}]\bar{w} =_\beta I$.

$\Leftarrow$: En effet, si $\bar{x} = Vlib(t)$, il est facile de vérifier que $t[\bar{y} := \bar{u}] = t[\bar{x} := \bar{u}']$, avec $u'_i = x_i$ si x_i n'est pas dans $\bar{y}$ et $u'_i = u_j$ si $x_i = y_j$. Il ne reste plus qu'à remarquer que $t[\bar{x} := \bar{u}'] = (\lambda\bar{x}.t)\bar{u}' =_{def} t^c\bar{u}'$.

Corollaire 20 *tw résoluble implique t résoluble.*

Corollaire 21 *La classe des termes résolubles est close par β-équivalence.*
Il en est donc de même de celle des termes non résolubles.

Preuve. Découle immédiatement de la proposition précédente et du fait que la β-équivalence est substitutive et contextuelle.

Lemme 22 *Tout terme β-équivalent à une fnt est résoluble.*

Preuve. Comme la classe des termes résolubles est close par β-équivalence il suffit de montrer la proposition pour les fnt. Comme la clôture d'une fnt est encore une fnt, il suffit de montrer l'implication pour les fnt closes. Si $t = \lambda\bar{x}.y\bar{u}$ est un tel terme, alors nécessairement $y \in \bar{x}$. Soit $y \equiv x_j$. Soit $\bar{z}$ une suite de variables de même longueur que $\bar{u}$ et soit $\bar{v}$ une suite de termes de même longueur que $\bar{x}$ et telle que $v_j = \lambda\bar{z}.I$. Alors $t\bar{v} \rightarrow_\beta I$.

Corollaire 23 *Tout terme normalisable ou h-normalisable est résoluble.*

Thorme 24 ***de normalisation de tête.*** *résoluble$\Longleftrightarrow$h-normalisable*

Preuve. La direction $\Leftarrow$ vient d'être prouvée, la direction $\Rightarrow$ sera prouvée plus tard.

L'équivalence entre les deux propriétés est d'autant plus intéressantes que certaines propriétés des termes peuvent être très faciles à montrer à partir d'une des définitions et beaucoup moins, ou pas du tout, avec l'autre. Par exemple, les résultats suivants sont des conséquences immédiates du Théorème de normalisation de tête et de lemmes précédemment montrés.

Corollaire 25 *La classe des termes normalisables de tête est close par β-équivalence.*

Corollaire 26 *Pour tous x et t on a : t résoluble ssi tx est résoluble ssi $\lambda x.t$ est résoluble ssi $\lambda x.tx$ est résoluble.*

C'est un exercice intéressant que d'essayer de démontrer ces résultats sans le Théorème de normalisation de tête.

3 Liens entre les théorèmes de normalisation.

Proposition 27 *Le Théorème de normalisation de tête implique le Théorème de normalisation gauche.*

Esquisse de preuve. Tout d'abord on remarque que tous les termes normalisables étant résolubles (Corollaire 23), ils sont normalisables de tête si on admet le théorème de normalisation de tête.

Considérons maintenant un terme normalisable u, sa forme normale n et sa forme normale de tête $\lambda\bar{x}.y\overline{u}$. La preuve se fait par induction sur la longueur de n. Par Church-Rosser il existe une réduction de la forme $u \to_h \lambda\bar{x}.y\overline{u} \to_\beta n$. Nécessairement n est donc de la forme $\lambda\bar{x}.y\bar{n}$, avec $l(\bar{n}) = l(\bar{u})$, chaque n_i étant la forme normale de u_i. Comme les n_i sont tous de longueur strictement inférieure à n, on peut appliquer l'hypothèse d'induction et en déduire que la réduction gauche de chaque u_i termine (nécessairement sur n_i). Comme la réduction gauche de u commence par la réduction de tête de u et se continue par la réduction gauche de u_1, puis u_2, etc., et que toutes ces réductions partielles terminent, la réduction gauche de u termine aussi. cqfd.

Proposition 28 *Le Théorème de normalisation gauche implique le Théorème de normalisation de tête.*

Preuve. Il suffit de le voir pour les termes clos. Soit donc t un terme résoluble clos et $\bar{v}$ tel que $t\bar{v} \to_\beta I$. Comme $t\bar{v}$ est normalisable il est g-normalisable, et donc h-normalisable. Soit $\bar{x}$ une suite de variables fraîches, de même longueur que $\bar{v}$. D'après le Corollaire 15 $t\bar{x}$ est h-normalisable. D'après le Lemme 11, t est donc h-normalisable.

4 Standardisation.

Dans cette section nous énonçons un Théorème, que nous appellerons de "pré-standardisation". Ce théorème exprime que si $t \to_\beta u$ alors on peut passer de t à u par une réduction de tête (partielle) suivie d'une réduction interne. Nous n'en démontrerons que la partie la plus facile (le cas où t est normalisable de tête), mais nous montrerons qu'il implique un Théorème de Standardisation sur les réductions, qui lui-même implique trivialement le Théorème de normalisation gauche.

Dfinition 29 *On appelle* redex interne *d'un terme t tout autre redex que le redex de tête.* Une réduction interne *est par définition une suite $t_1...t_n...$ de termes tels que t_{i+1} est obtenu en contractant un redex interne de t.*

En d'autres termes une réduction interne est une suite de réductions qui ne touche pas au "$(\lambda''$ de tête. En se paraphrasant une nouvelle fois : si le terme s'écrit $\lambda\bar{x}.v\bar{t}$, avec v variable ou v redex avec $v \equiv (\lambda x.u)w$, alors une réduction interne réduit dans les t_i et éventuellement dans u et w. En particulier une réduction interne ne créera pas de λ de tête (n'allongera pas $\bar{x}$).

Notation 30 *Soit σ une β-réduction, on notera $l(\sigma)$ sa longueur (au sens du nombre d'étapes de β_0 réductions), qui est un élément de $N \cup \{\infty\}$. Si σ est finie et est une suite de réduction de t, on notera $\sigma(t)$ le terme auquel elle aboutit. Si σ, τ sont deux suites finies de réduction telles que τ parte du terme où σ aboutit, alors on notera $\sigma\tau$ la suite obtenue en effectuant σ suivie de τ.*

Lemme 31 *Soient t et u tels que $t \to_\beta u$. Alors :*
(i) ou bien toute réduction de t sur u est interne
(c'est le cas par exemple si t est une fnt)
(ii) ou bien il y a une réduction $\sigma : t \to_\beta u$ telle que σ_1 réduit le rédex de tête de t.

Preuve. Soit τ une réduction non interne de t sur u, m la longueur de τ et $k \leq m$ le premier entier tel que τ_k réduit le redex de tête de t_k. Si $t \equiv \lambda\bar{y}.(\lambda x.u)v\bar{w}$ on a $t_k = \lambda\bar{y}.(\lambda x.u')v'\bar{w}'$, où r' dénote un β-réduit de r $(r = u, v, \bar{w})$. On prend alors $\sigma = \sigma_1\sigma'\tau_{k+1}...\tau_m$ avec :
$\sigma_1 : t \equiv \lambda\bar{y}.(\lambda x.u)v\bar{w} \to_{\beta_0} \lambda\bar{y}.u[x := v]\bar{w} =_{def} t$" et
$\sigma' :$ t" $=_{def} \lambda\bar{y}.u[x := v]\bar{w} \to_\beta \lambda\bar{y}.u'[x := v']\bar{w}' =_{def} t_{k+1}$.

Thorme 32 ***de pré-standardisation.***
Pour tous t, u tels que $t \to_\beta u$ on peut passer de t à u par une réduction de tête σ^h suivie d'une réduction interne σ^i.

Preuve. Si t est h-normalisable c'est une conséquence directe du lemme (induction sur la longueur de la réduction de tête de t).

Il existe plusieurs preuves du cas général (ou plusieurs variantes de preuves). On en trouvera une dans [Bar], une autre dans [K, version anglaise], utilisant du

λ-calcul étiquetté, et enfin [Ron] contient une preuve syntaxique directe (d'une page) utilisant la réduction parallèle $\Rightarrow$ dont nous avons déja parlé.

Ce théorème est intéressant à connaître, mais nous ne l'utiliserons que dans ce chapitre.

Corollaire 33 *Si $t \to_\beta u$ et si u est une fnt, alors t est normalisable de tête et il existe une réduction de t sur u qui commence par la réduction de tête complète de t.*

Preuve. Soient σ^h et σ^i les réductions données par le Théorème précédent. Comme σ^i est interne et u est une fnt, nécessairement $\sigma^h(t)$ est une fnt. qed.

Corollaire 34 *Tout terme β-équivalent à une abstraction a un réduit de tête qui est une abstraction.*

Preuve. En effet si $t' =_\beta t$ et $t \to_h \lambda x.u$. En utilisant Church-Rosser on déduit qu'il existe une β-réduction σ de t' sur une abstraction u'. Soit $\sigma' =_{def} \sigma^h \sigma^i$ une réduction obtenue en appliquant le théorème de pré-standardisation. Comme $\sigma'(t')$ est une abstraction et que σ^i ne crée pas de λ de tête, nécessairement $\sigma^h(t)$ est une abstraction. q.e.d.

Dfinition 35 *Etant donnée une suite de réductions σ, on définit $d(k)$ comme le nombre de symboles qui précédent le "$(\lambda$" qui est réduit à l'étape k (les réduits successifs étant supposés écrits en notation formelle complète). Une β-réduction (finie ou infinie) est* standard *si pour tout $k < l(\sigma)$ on a $d(k) \leqslant d(k+1)$.*

Intuitivement : on ne revient jamais en arrière.

Exemple 36 *La réduction de tête et la réduction gauche sont des réductions standards.*

Exemple 37 *La réduction $\delta\underline{(\delta_3)(\delta)I} \to_{\beta 0} \delta((\delta)I)((\delta)I)\underline{(\delta)I} \to_{\beta 0} \delta((\delta)I)((\delta)I)\underline{(I)I} \to_{\beta 0}$ $\delta((\delta)I)((\delta)I)I$ est standard (à chaque étape on a souligné le redex qui est réduit), et on ne peut pas continuer à réduire en restant standard.*

Remarquons que dans l'exemple ci-dessus nous n'avons utilisé tout de même une écriture un peu abbrégée; mais il est sous-entendu que c'est l'écriture formelle complète qui fait référence.

Thorme 38 ***de standardisation**. Pour tous termes t, u tels que $t \to_\beta u$ on peut passer de t à u par une réduction standard.*

Preuve. Par induction sur $l(u)$. Soient σ^h et σ^i les réductions données par le Théorème précédent. Soient $\bar{x}, v, \bar{t}$ telle que $\sigma^h(t) = \lambda\bar{x}.v\bar{t}$ avec v *variable ou redex.* Comme σ^i est interne on a nécessairement $u = \lambda\bar{x}.v'\bar{t}'$ avec $v \to_\beta v'$ et $t_i \to_\beta t'_i$ pour chaque i. Comme v' et les t'_i sont de longueur strictement inférieure à u on peut appliquer l'hypothèse d'induction et remplacer chacune de ces réductions par une réduction standard; soient τ_v et τ_i les réductions standard correspondantes de $\sigma^h(t)$. Alors la réduction $\sigma^h\tau_v\tau_1...\tau_{l(\bar{t})}$ est une réduction standard de t sur u.

Corollaire 39 ***Théorème de normalisation gauche.***

Preuve. On ne peut arriver sur un terme normal par une réduction standard que si c'est la réduction gauche.

5 Arbres de Böhm et de Lévy-Longo.

L'arbre de Böhm $BT(t)$ d'un terme t représente essentiellement l'information maximale que l'on peut glaner sur t au cours de toutes ses possibles réductions, en considérant uniquement les couples $(\lambda\bar{x}.y, n)$ comme des informations valables ; *l'arbre de Lévy-Longo* $LL(t)$ est une variante qui considère aussi comme information l'ordre de fonctionnalité des termes non h-normalisables (ce qui est utile en cas de stratégie paresseuse).

Dans les deux cas il s'agit d'arbres qui ont une racine (placée en haut !) et des branches, finies ou non, qui descendent vers le bas, et qui branchent finiment à chaque noeud qui n'est pas une feuille (ces arbres peuvent se définir formellement comme des graphes particuliers, ou comme des ensembles particuliers de suites d'entiers (voir par exemple [Bar])), de plus ces arbres sont étiquetés (chaque noeud porte une information, ou l'étiquette $\perp$ qui correspond a “information nulle”).

On définit $BT(t)$ de la manière suivante [4].

Si t n'est pas normalisable de tête, alors $BT(t) = \perp$, en ce sens que $BT(t)$ est réduit à sa racine (qui est alors aussi une feuille), étiquettée $\perp$.

Si $t \to_h \lambda\bar{x}.y\bar{t}$, alors $BT(t)$ a une racine étiquettée $\lambda\bar{x}.y$ d'où partent les n sous-arbres $BT(t_1)$, ..., $BT(t_n)$ (de gauche à droite) ; en particulier si $n = 0$ l'arbre est réduit à sa racine (il existe une définition alternative, par induction sur la profondeur des noeuds (cf. [Bar]). Comme nous ne ferons pas dans ce cours de démonstration sur les arbres, il est inutile que nous donnions cette définition analytique.

Le rapport avec les discussions précédentes est le suivant :

Si t est normalisable alors $BT(t)$ est fini et est juste une description arborescente de la forme normale de t. De plus la réduction gauche de t construira l'arbre de Böhm de gauche à droite (en commençant par construire intégralement la branche la plus à gauche, puis etc. [A voir sur un exemple !]).

Si t n'est pas normalisable alors $BT(t)$ peut-être infini ou fini, dans ce dernier cas il contient nécessairement au moins un “$\perp$”.

Comme exemple de terme t tel que $BT(t)$ est infini on peut considérer $t =_{def} Yf$, où f est une variable : son arbre de Böhm consiste en une seule branche infinie, dont tous les noeuds sont étiquettés “f”.

$LL(t)$ est défini de manière analogue sauf que si t n'est pas résoluble on regarde le nombre maximal $k(t)$ de λ qui peuvent arriver en tête, lors de la

[4] Il s'agit en fait d'une définition rigoureuse “par co-induction”, mais je n'insisterai pas sur ce point.

réduction de tête de t (nous verrons plus tard que c'est en fait son ordre de fonctionnalité) ; si $k(t) = m \geqslant 0$ alors on étiquette par $\lambda x_1 ... \lambda x_m . \perp$ et si $k(t)$ est infini on étiquette par "$\top$".

Il existe une classe de termes intéressante qui est intermédiaire entre les termes normalisables et les h-normalisables : il s'agit des termes héréditairement h-normalisables, qui sont aussi, modulo le théorème de normalisation de tête, les termes héréditairement résolubles.

Dfinition 40 *Un terme t est* héréditairement normalisable de tête *si il est normalisable de tête et si $t \rightarrow_h \lambda \bar{x}. y \bar{t}$ avec tous les t_i héréditairement normalisables de tête* [5]*.*

Un exemple est Y (ou Yf, $f \in V$) et on peut les définir alternativement comme étant les termes t tels que $BT(t)$ ne contient pas $\perp$. Par suite :

Remarque 41 *Si t est héréditairement résoluble (en particulier si t est normalisable), alors $BT(t) = LL(t)$. La réciproque est évidemment fausse.*

Remarque 42 *Il est clair que la relation* "avoir même arbre de Böhm" *est une relation d'équivalence qui contient la β-équivalence ; en fait c'est une λ-congruence mais le fait qu'elle est contextuelle n'est pas trivial du tout (voir [Bar]). C'est évidemment le cas de l'application qui fait problème.*

Remarque 43 *Il existe des versions de la notion d'arbre de Böhm qui engendrent des $\lambda\eta$-congruences.*

[5]Il s'agit encore d'une définition par co-induction.