λ-calcul

Syntaxe :

Définition. Soit un ensemble $\{x, y, z, \dots\}$ dénombrable de variables. On définit les *termes de λ-calcul* (ou *λ-termes*) par induction :

- les variables sont des termes ;
- si M et N sont des termes, alors (MN) est un terme ;
- si M est un terme et x une variable, alors $(\lambda x.M)$ est un terme.

On dit que M et N sont des *sous-termes* de (MN) et que M est un *sous-terme* de $(\lambda x.M)$

Exemple : les expressions $(\lambda x.(xy))$ et $((\lambda y.y)(\lambda x.(xy)))$ sont des termes. L'expression $(x.\lambda x)$ n'est pas un terme.

On s'autorisera à écrire MNP pour $((MN)P)$ où M, N et P sont des termes et $\lambda xy.M$ pour $(\lambda x.(\lambda y.M))$ où M est un terme et x, y des variables.

Définition. Dans un terme $\lambda x.M$, on dit que M et tous les sous-termes de M sont dans la *portée* de λx.

Définition. Dans un terme P, une occurrence x est

- *liée* si x est un sous-terme de P et que x est dans la portée d'un λx qui est dans P.
- *libre* si elle n'est pas liée.

Définition. Si un terme n'a pas de variable libre, il est dit *clos*.

Exemple : les termes $I = \lambda x.x$, $K = \lambda xy.x$ et $S = \lambda xyz.xz(yz)$ sont clos.

Définition. Soient M, N des termes et x une variable. La *substitution* de x par N dans M, notée $M[x \to N]$, est le remplacement de chaque occurrence libre de x dans M par N en renommant si besoin les variables de M par des variables fraîches pour qu'elles soient libres dans $M[x \to N]$.

Exemple : l'expression $(\lambda y.xy)[x \to \lambda y.y]$ nécessite un renommage, $(\lambda z.(\lambda y.y)z)$ est alors l'expression obtenue.

Définition. On définit l'*α-conversion* comme la plus petite relation d'équivalence $\sim_\alpha$ satisfaisant :

- $\lambda x.M \sim_\alpha \lambda y.\,(M[x \to y])$ pour une variable fraîche y ;
- $\lambda x.M \sim_\alpha \lambda x.N$ si $M \sim_\alpha N$;
- $MN \sim_\alpha PQ$ si $M \sim_\alpha P$ et $N \sim_\alpha Q$.

Exemple : on a $\lambda xy.x(xy) \sim_\alpha \lambda uv.u(uv)$

β-réduction :

Définition. On définit la relation de *β-réduction*, notée $\to_\beta$, par induction sur les termes

- $(\lambda x.M)N \to_\beta M[x \to N]$;
- si $M \to_\beta M'$, alors $\lambda x.M \to_\beta \lambda x.M'$;
- si $M \to_\beta M'$, alors $MN \to_\beta M'N$;
- si $N \to_\beta N'$, alors $MN \to_\beta MN'$.

Exemples :

- si $I = \lambda x.x$ alors $IM \to_\beta M$;
- si $K = \lambda xy.x$ alors $(KM)N \to_\beta (\lambda y.M)N \to_\beta M$;
- si $S = \lambda xyz.xz(yz)$ alors $SMNP \to_\beta^3 MP(NP)$.

Définition. Un terme M est sous *forme normale* s'il ne peut pas être β-réduit.

Lemme.

- une variable x est sous forme normale ;
- si M est sous forme normale, alors $\lambda x.M$ est sous forme normale.

Lemme. Si $P \to_\beta^* P'$ et $Q \to_\beta^* Q'$ alors $Q[x \to P] \to_\beta^* Q'[x \to P']$.

Théorème.(Church-Rosser I) Si $P \to_\beta^* P_1$ et $P \to_\beta^* P_2$ alors il existe Q tel que $P_1 \to_\beta^* Q$ et $P_2 \to_\beta^* Q$. Cette propriété est nommée *confluence*.

Définition. On dit que P possède une forme normale s'il existe un terme M sous forme normale tel que $P \to_\beta^* M$. Par le théorème de Church–Rosser, si P possède une forme normale alors elle est unique.

Définition. On dit que M et N sont *β-équivalents*, noté $M \sim_\beta N$, s'il existe $M_1, M_2, \ldots, M_n$ tels que $M_1 \sim_\alpha M$, $M_n \sim_\alpha N$ et pour tout $i \in [\![1; n-1]\!]$, on a $M_i \to_\beta^* M_{i+1}$ ou $M_{i+1} \to_\beta^* M_i$.

Théorème.(Church-Rosser II) Si $P \sim_\beta Q$ alors il existe T tel que $P \to_\beta^* T$ et $Q \to_\beta^* T$.

Corollaire.

- Si $P \sim_\beta Q$ et si Q est sous forme normale alors $P \to_\beta^* Q$.
- Si $P \sim_\beta Q$ alors soit ils ont la même forme normale, soit ils n'en ont pas.

Exemples :

On va maintenant définir différents termes de λ-calcul.

- $T := \lambda xy.x$
- $F := \lambda xy.y$
- $\mathsf{if\ then\ else} := \lambda bxy.bxy$

La terminologie est justifée par le fait que si $X \sim_\beta T$, alors $\mathsf{if}\ X\ \mathsf{then}\ M\ \mathsf{else}\ N \sim_\beta T\ M\ N \sim_\beta M$, de même si $X \sim_\beta F$, alors $\mathsf{if}\ X\ \mathsf{then}\ M\ \mathsf{else}\ N \sim_\beta F\ M\ N \sim_\beta N$.

- $\langle M, N\rangle := \lambda x.(xMN)$
- $\mathsf{fst} := T$
- $\mathsf{snd} := F$

On a $\langle M, N\rangle \mathsf{fst} \sim_\beta T\ M\ N \sim_\beta M$

- $[M] := M$ (liste contenant uniquement l'élément M)
- $[M_1, \ldots, M_{n+1}] := \langle M_1, [M_2, \ldots, M_{n+1}]\rangle$
- $\pi_i^n := \lambda x_1 \cdots x_n.x_i$ (projection)

On note $M^n N = \begin{cases} N & \text{si } n = 0, \\ M(M^{n-1}N) & \text{si } n > 0. \end{cases}$

Représentation des entiers :

On veut pouvoir représenter les entiers et utiliser les opérations usuelles dessus.

Entiers de *Church* :

- $[\![0]\!]_C := \lambda sz.z$
- $[\![n]\!]_C := \lambda sz.s^n z$
- $[\![\mathsf{succ}]\!]_C := \lambda n.(\lambda sz.ns(sz))$

La terminologie est justifée par le fait que

$$\begin{aligned}[\![\mathsf{succ}]\!]_C[\![n]\!]_C &\sim_\beta \lambda sz.[\![n]\!]_C s(sz) \sim_\alpha \lambda sz.(\lambda uv.u^n v)s(sz)\\ &\sim_\beta \lambda sz.s^n(sz) \sim_\alpha \lambda sz.s^{n+1}z \sim_\alpha [\![n+1]\!]_C.\end{aligned}$$

- $[\![+]\!]_C := \lambda n_1 n_2.\lambda sz.n_1 s(n_2 sz) = \lambda n_1 n_2.n_1[\![succ]\!]_C n_2$
- $[\![\mathsf{exp}]\!]_C := \lambda mn.mn$
- $[\![\mathsf{pred}]\!]_C := \lambda n.\mathsf{fst}(n(\lambda y.[\mathsf{snd}y, [\![\mathsf{succ}]\!]_C\mathsf{snd}y][[\![0]\!]_C, [\![0]\!]_C]))$

On incrémente n fois le couple $(k, k+1)$ jusqu'à $(n-1, n)$ puis on extrait la première coordonnée afin d'obtenir $n-1$.

- $[\![\mathsf{is\_zero}]\!]_C := \lambda n.n(\lambda x.F)T$

La terminologie est justifée par le fait que

$$\begin{aligned}[\![\mathsf{is\_zero}]\!]_C[\![0]\!]_C &\sim_\beta [\![0]\!]_C(\lambda x.F)\ T\\ &\sim_\alpha (\lambda sz.z)(\lambda x.F)\ T \sim_\beta T\\ [\![\mathsf{is\_zero}]\!]_C[\![n+1]\!]_C &\sim_\beta [\![n+1]\!]_C(\lambda x.F)\ T \sim_\alpha (\lambda sz.s^{n+1}z)(\lambda x.F)\ T\\ &\sim_\beta (\lambda x.F)^{n+1}\ T \sim_\alpha (\lambda x.F)^n((\lambda x.F)\ T)\\ &\sim_\beta (\lambda x.F)^n\ F \sim_\beta F\end{aligned}$$

Entiers de *Barendregt*

- $[\![0]\!]_B := I = \lambda x.x$
- $[\![n+1]\!]_B := \langle F, [\![n]\!]_B\rangle = \lambda x.(xF[\![n]\!]_B)$
- $[\![\mathsf{succ}]\!]_B := \lambda n.\langle F, n\rangle$
- $[\![\mathsf{pred}]\!]_B := \lambda n.nF$

La terminologie est justifée par le fait que

$$\begin{aligned}[\![\mathsf{pred}]\!]_B[\![n+1]\!]_B &\sim_\alpha \lambda n.(nF)[\![n+1]\!]_B \sim_\beta [\![n+1]\!]_B\ F \sim_\alpha \lambda x.(xF[\![n]\!]_B)\ F\\ &\sim_\beta F\ F\ [\![n]\!]_B \sim_\alpha (\lambda xy.y)F\ [\![n]\!]_B \sim_\beta [\![n]\!]_B.\end{aligned}$$

- $[\![\mathsf{is\_zero}]\!]_B := \lambda n.nT$

La terminologie est justifée par le fait que

$$\begin{aligned}[\![\mathsf{is\_zero}]\!]_B[\![0]\!]_B &\sim_\alpha (\lambda n.nT)[\![0]\!]_B \sim_\beta [\![0]\!]_B\ T\\ &\sim_\alpha (\lambda x.x)\ T \sim_\beta T\\ [\![\mathsf{is\_zero}]\!]_B[\![n+1]\!]_B &\sim_\alpha (\lambda n.nT)[\![n+1]\!]_B \sim_\beta [\![n+1]\!]_B\ T\\ &\sim_\alpha \lambda x.(x\ F\ [\![n]\!]_B)\ T \sim_\beta T\ F\ [\![n]\!]_B \sim_\beta F\end{aligned}$$

Point fixe :

Définition. Un *point fixe* est un λ-terme Y tel que, pour tout terme G, on a

$$YG \sim_\beta G(YG).$$

Exemple : le point fixe de Turing est $\theta = AA$ avec $A = \lambda xf.f(xxf)$
En effet, pour tout terme G, on a

$$\begin{aligned}\theta G &\sim_\alpha AAG \\ &\sim_\alpha (\lambda xf.f(xxf))(\lambda xf.f(xxf))G \\ &\to_\beta (\lambda f.f((\lambda xz.z(xxz))(\lambda xz.z(xxz))f))G \\ &\to_\beta G((\lambda xz.z(xxz))(\lambda xz.z(xxz))G) \\ &\sim_\alpha G(AAG) \\ &\sim_\alpha G(\theta G).\end{aligned}$$

Exemple : le point fixe de Curry est $Y = \lambda f.((\lambda x.f(xx))(\lambda x.f(xx)))$
En effet, pour tout terme G, on a

$$\begin{aligned}YG &\to_\beta (\lambda x.G(xx))(\lambda x.G(xx)) \\ &\to_\beta G((\lambda x.G(xx))(\lambda x.G(xx))) \\ &\leftarrow_\beta G(YG).\end{aligned}$$

λ-calcul et calculabilité :

Définition. Soit ϕ une fonction de $\mathbb{N}^p$ dans $\mathbb{N}$. La fonction ϕ est λ-*définissable* si et seulement s'il existe un λ-terme G tel que

$$\forall n_1, \dots, n_p \in \mathbb{N}, \quad G[\![n_1]\!] \dots [\![n_p]\!] \sim_\beta [\![\phi(n_1, \dots, n_p)]\!]$$

Théorème. Une fonction est μ-récursive si et seulement si elle est λ-définissable.

(voir μ-recursif $\implies$ λ-définissable)

On dit que deux ensembles $\mathcal{A} \subset \mathbb{N}$ et $\mathcal{B} \subset \mathbb{N}$ disjoints sont *récursivement séparables* s'il existe une fonction μ-récursive φ telle que

$$x \in \mathcal{A} \implies \varphi(x) = 0 \quad \text{et} \quad x \in \mathcal{B} \implies \varphi(x) = 1.$$

Après avoir transposé la définition de *récursivement séparable* pour les λ-termes, on peut démontrer le théorème suivant.
Théorème. Aucune paire d'ensembles non vides de termes clos par β-équivalence ne sont récursivement séparables.

(voir Théorème de Scott–Curry)