THEORIE DES COMBINATEURS
(Présentation formelle)

L'INDUCTION

Patrick Bellot
Télécom ParisTech

Théorie formelle

- Nous allons présenter la *théorie formelle* des combinateurs.
- Une théorie formelle est une théorie ou axiomes et règles de déductions n'utilisent que *la forme* des « objets » dont elle parle, c'est-à-dire *la syntaxe*.
- Cela n'empêche nullement la théorie d'avoir *un sens*, c'est-à-dire *une interprétation* en termes d'objets compréhensibles (les fonctions processus pour la théorie des combinateurs mais ce n'est pas la seule interprétation possible).
- La notion d'*interprétation* est aussi mathématisée.
- Toute théorie formelle se présente à la manière de la théorie des combinateurs. *C'est aussi le cas des systèmes logiques.*

Définition par induction (structurelle)

- Si l'on veut définir un ensemble par induction structurelle, on utilise :
  - une ou plusieurs *règles de base* établissant que certains « objets » sont éléments de cet ensemble ;
  - une ou plusieurs règles de *pas d'induction* permettant de construire de nouveaux éléments à partir d'élements déjà construits.
- Exemple :
  - 0 est un entier naturel (BASE)
  - si n est un entier naturel, n+1 est un entier naturel (PAS)

Définition par induction (structurelle)

- Toute définition par induction structurelle est implicitement (ou explicitement) suivie de la *règle de fermeture* qui dit :
  - tout objet de l'ensemble est construit à l'aide des règles précédentes appliquées un nombre fini de fois.

- Sans la règle de fermeture, la définition est ambiguë car elle n'exclut pas l'application un nombre infini de fois d'une règle de pas d'induction ET elle ne dit pas ce qui n'est pas un élément de l'ensemble…

Définition par induction (structurelle)

- Toute définition de structure chaînée en informatique, liste ou arbre par exemple, doit être présentée par induction.

  L'implémentation n'est qu'un moyen de réaliser informatiquement cette structure de données.

- Des langages tels que ML (CAML pour les *froggies*) prennent en compte les structures de données définies par induction.

Alphabet de la théorie des combinateurs

Une théorie formelle utilise un langage et donc un alphabet qui doit être un ensemble au plus dénombrable de symboles ayant des fonctions spécifiques. Dans le cas des combinateurs :

- les lettres S et K servent à désigner des *constantes* appelés *combinateurs* ;
- les lettres minuscules, a,b,c,… , éventuellement indicées par des entiers servent à désigner des *variables* ;
- les parenthèses (et) sont utilisées pour construire des objets appelés *applications* ;
- := est le symbole de *réduction*, = celui d'*égalité*.

Les termes de la théorie des combinateurs

- Les termes de la théorie des combinateurs sont définis par induction structurelle :
  - S et K sont des termes ; (BASE)
  - toute variable est un terme ; (BASE)
  - si P et Q sont des termes, (P Q) est un terme ; (PAS)
  - règle de fermeture.

- S, K et les variables sont appelés des atomes.
- On adopte par la suite la convention d'associativité à gauche de l'application qui permet d'éliminer les parenthèses superflues mais ne fait pas partie de la théorie.

Intérêt de l'induction

- Comme les termes sont définis par induction structurelle, on peut définir des « notions à propos des termes » par *induction sur la structure des termes*. Ce sont des définitions qui suivent la définition des termes.

- Exemple : les sous-termes d'un terme.
  - on donne explicitement la définition des sous-termes d'un terme pour chacun des éléments de la base ;
  - on donne le moyen de construire les sous-termes d'un terme (P Q) à partir des sous-termes de P et de ceux de Q.

Définition par induction sur le structure des termes

- Les sous-termes d'un terme :
  - si M est un atome ou une variable, son seul sous-terme est lui-même ; (BASE)
  - si M est une application (P Q), les sous-termes de M sont M, les sous-termes de P et les sous-termes de Q. (PAS)

- En abrégé :
  - si M est un atome, $ST(M) = \{ M \}$
  - Si $M \equiv (P\ Q)$, $ST(M) = \{M\} \cup ST(P) \cup ST(Q)$

Intérêt de l'induction

- Quand des objets sont définis par induction structurelle, on peut démontrer par induction une propriété de ces objets : démontrer que les objets de la base ont cette propriété et que la propriété respecte les pas d'induction.
- Exemple : les entiers naturels sont définis par :
  - 0 est un entier naturel ;
  - si n est un entier naturel, n+1 est un entier naturel

  Démontrer une propriété P(n) des entiers naturels :
  - démontrer P(0) ;
  - démontrer P(n+1) en supposant P(n).

Démonstration par induction pour les termes

- Pour démontrer une propriété P(M) des termes de la théorie des combinateurs, je dois :
  - démontrer P(K) et P(S) ; (BASE)
  - démontrer P(x) lorsque x est une variable ; (BASE)
  - démontrer P(M N) en supposant P(M) et P(N) (*hypothèse d'induction*). (PAS)

Exemple

- Propriété : tout terme peut s'écrire sous la forme a M_1 … M_n pour un $n \geq 0$ et a atomique. Le cas $n=0$ étant par commodité le cas ou M s'écrit a.
- Démonstration :
  - si M est un atome, constante ou variable, M est déjà sous la forme demandée avec $n=0$;
  - Si $M \equiv P\ Q$, on suppose par hypothèse d'induction que P possède la propriété donc P peut s'écrire a P_1 … P_u. Donc : $M \equiv P\ Q \equiv (a\ P_1 \ldots P_u)\ Q \equiv a\ P_1 \ldots P_u\ Q$ par application de la règle d'associativité à gauche de l'application.

L’induction

- L’induction est un outil mathématique très utilisé en logique et en informatique.
- Maîtriser l’induction est essentiel dès que l’on veut traiter de la logique ou de l’informatique.

Démontrer des formules

- Une théorie formelle a pour fonction de permettre de démontrer des formules.

- On ne parle pas de formules *vraies* ou *fausses* (ce qui est une notion discutée) mais de formules *démontrables* ou *non démontrables* dans le cadre du système formel avec les axiomes et les règles que l'on se donne (ce qui n'est pas discutable).

- Le choix des axiomes et des règles peut être discuté.

Les formules

- Les formules de la théorie des combinateurs sont les expressions de la forme M := N ou de la forme M = N avec M et N des termes.
- Une formule de la forme M := N se lit M *se réduit à* N.
- Une formule de la forme M = N se lit M *est égal à* N.

L'axiomatisation

- L'axiomatisation est l'ensemble des règles que l'on peut utiliser pour *démontrer* une formule.
- L'axiomatisation d'une théorie formelle comprend principalement deux types de règles :
  - des *axiomes* : ce sont des règles qui affirment que certaines formules sont admises comme étant démontrées ;
  - des *règles d'inférence* ou *règles de déduction* qui permettent de démontrer une nouvelle formule à partir de formules déjà démontrées, on les présente sous la forme :

$$\frac{H_1 \ldots H_k}{C}$$

$H_1 \ldots H_k$ ⟵ *hypothèses*

C ⟵ *conclusion*

Axiomes de la théorie des combinateurs

- Trois axiomes :

  (S) S X Y Z := X Z (Y Z)

  (K) K X Y := X

  (Id) X := X

- Un axiome comme (K) signifie que toute formule de la forme K X Y := X est considérée comme démontrée a priori.
- En fait, il dénote une infinité (dénombrable) d'axiomes, un pour chacune des valeurs possibles des méta-variables X et Y.
- Un tel type d'axiome est appelé un *schéma d'axiome*.

Règles de la théorie des combinateurs

- Trois règles pour la réduction :

$$\frac{M := N \quad N := L}{M := L}\ (tr)$$

$$\frac{M := N}{M\ L := N\ L}\ (al)$$

$$\frac{M := N}{L\ M := L\ N}\ (ar)$$

Règles de la théorie des combinateurs

- Trois règles pour l'égalité : l'égalité est la fermeture symétrique et transitive de la réduction !

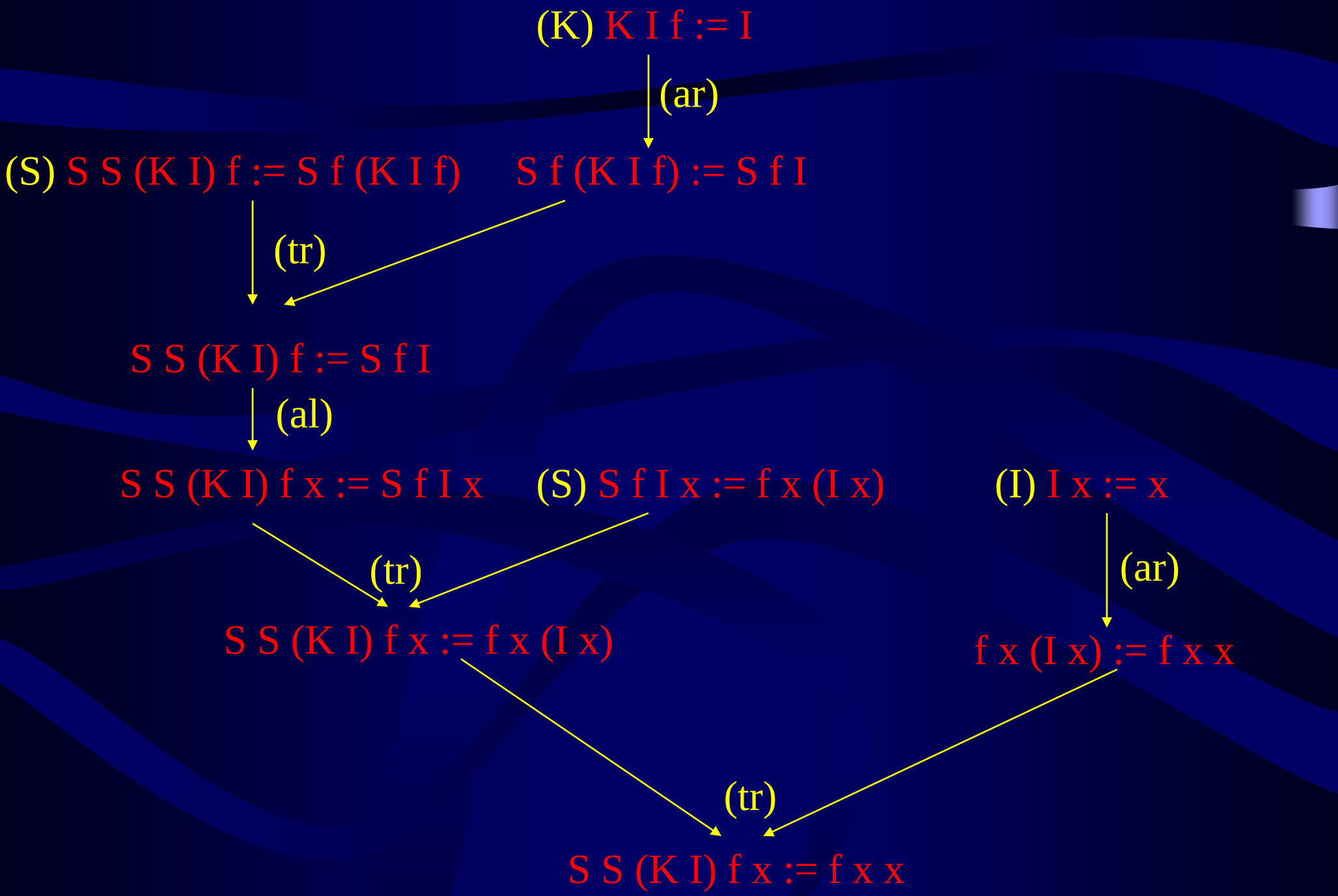
$$\frac{M := N}{M = N}\ (fe)$$

$$\frac{M = N}{N = M}\ (rf)$$

$$\frac{M = N \quad N = L}{M = L}\ (te)$$

Démonstration

- Une démonstration dans un tel système prend alors la forme d'un arbre :
  - la racine de l'arbre est la formule démontrée ;
  - une feuille de l'arbre est un axiome ;
  - on passe des fils d'un nœud de l'arbre au nœud par l'application d'une règle d'inférence.
- Exemple suit !


(K) K I f := I
(ar)
(S) S S (K I) f := S f (K I f)
S f (K I f) := S f I
(tr)
S S (K I) f := S f I
(al)
S S (K I) f x := S f I x
(S) S f I x := f x (I x)
(I) I x := x
(tr)
(ar)
S S (K I) f x := f x (I x)
f x (I x) := f x x
(tr)
S S (K I) f x := f x x

(K) K I f := I
———————————— (ar)
(S) S S (K I) f := S f (K I f)   S f (K I f) := S f I
———————————————————————————— (tr)
S S (K I) f := S f I
———————————— (al)
S S (K I) f x := S f I x   (S) S f I x := f x (I x)
———————————————————————————— (tr)
(I) I x := x
———————————— (ar)
S S (K I) f x := f x (I x)   f x (I x) := f x x
———————————————————————————— (tr)
S S (K I) f x := f x x

Cette présentation plus « littéraire » et plus concise
a la faveur des ouvrages de logique.

Démonstrations

- Dans les deux cas, nous avons rigoureusement et formellement démontré que la formule

  S S (K I) f x := f x x

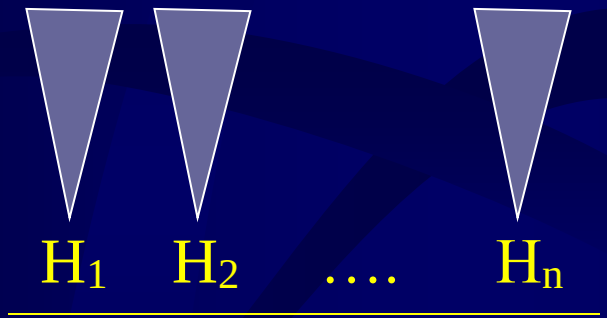
  était un théorème de notre système formel.

- L'arbre EST la preuve.

- Nous l'avions déjà démontré INFORMELLEMENT :
  S S (K I) f x := S f (K I f) x := S f I x := F x (I x) := f x x

Intérêt des preuves sous forme d'arbres

- Les preuves-arbres permettent d'élégantes démonstrations par récurrence sur la *hauteur* de la preuve:
  - une preuve de hauteur 0 est un axiome ;
  - une preuve de hauteur k+1 est de la forme :



où les preuves de $H_1,H_2,..,H_k$ sont de hauteur au plus k.

Exemple

- Théorème : si M = N alors il existe une suite de termes $M_0,\ldots,M_k$ telle que :

$$M \equiv M_0 :=: M_1 :=: \ldots :=: M_k \equiv N$$

où :=: désigne soit := soit =:

- Plus clairement, si deux termes sont égaux, il existe une suite d'expansions et de réductions reliant le premier au deuxième.
- La démonstration se fait par récurrence sur la hauteur de l'arbre. Comme il n'existe pas d'axiome décrivant une égalité, il n'existe pas d'arbre de preuve de hauteur 0. La récurrence commence donc à 1.

Pour un arbre de hauteur 1

- Un arbre de hauteur 1 est un ou plusieurs axiomes suivi de l'application d'une règle qui fournit la conclusion M = N.
- Seules trois règles permettent une conclusion de la forme M=N : (fe), (rf), (te).
- Les hypothèses des règles (rf) et (te) ne peuvent être des axiomes puisque ce sont des égalités. La dernière règle est donc (fe).
- La preuve de hauteur 1 est donc :

$$\frac{M := N}{M = N}$$

où M := N est un axiome.

- Le résultat est alors évident.

$$\frac{M := N}{M = N}\ (fe) \qquad \frac{M = N}{N = M}\ (rf) \qquad \frac{M = N \quad N = L}{M = L}\ (te)$$

Pour un arbre de hauteur k+1

- On examine la dernière règle utilisée dans la preuve. C'est une règle dont la conclusion est de la forme M=N. Il ne peut s'agir que de l'une des règles (fe), (rf) ou (te).
- Si la dernière règle est (fe), le résultat est trivial.
- Si la dernière règle est (rf), l'hypothèse de récurrence s'applique à la formule N=M. Il existe donc une suite de réduction et d'expansions allant de N à M. En inversant cette suite, on en obtient une qui va de M à N.
- Si la dernière règle est (te), l'hypothèse de récurrence s'applique aux hypothèses M=L et L=N. Il existe donc deux suites d'expansions et de réductions, l'une entre M et L et l'autre entre L et N. En les mettant bout à bout, on en obtient une de M à N.

$$\frac{M := N}{M = N}\ (fe)$$

$$\frac{N = M}{M = N}\ (rf)$$

$$\frac{M = L \quad L = N}{M = N}\ (te)$$

La technique de l'examen de la dernière règle est un classique…

Le théorème de CHURCH-ROSSER

- Si $M=N$, il existe un terme Z tel que $M :=Z$ et $N:=Z$.

- Démonstration : si $M=N$, il existe une suite $M_0,\ldots,M_k$ telle que $M \equiv M_0 :=: M_1 :=: \ldots :=: M_k \equiv N$.

  Puis on démontre que si $k>2$, il existe une telle suite de longueur strictement inférieure à k.

  En itérant ce procédé, on trouve une suite de longueur au plus 2.

1. $M \equiv M_0 := M_1 := M_2 :=: M_3 :=: \ldots :=: M_k \equiv N$
   de $M_0 := M_1$ et $M_1 := M_2$, on déduit par (tr) que $M_0 := M_2$, donc:
   $M \equiv M_0 := M_2 :=: M_3 :=: \ldots :=: M_k \equiv N$
2. $M \equiv M_0 := M_1 =: M_2 := M_3 :=: \ldots :=: M_k \equiv N$
   de $M_1 =: M_2 := M_3$, on déduit par le lemme C-R $M_1 := Z =: M_3$, donc:
   $M \equiv M_0 := M_1 := Z =: M_3 :=: \ldots :=: M_k \equiv N$
   de $M_0 := M_1$ et $M_1 := Z$, on déduit par (tr) que $M_0 := Z$, donc:
   $M \equiv M_0 := Z =: M_3 :=: \ldots :=: M_k \equiv N$
3. $M \equiv M_0 := M_1 =: M_2 =: M_3 :=: \ldots :=: M_k \equiv N$
   de $M_1 =: M_2$ et $M_2 =: M_3$, on déduit par (tr) que $M_1 =: M_3$, donc:
   $M \equiv M_0 := M_1 =: M_3 :=: \ldots :=: M_k \equiv N$
4. $M \equiv M_0 =: M_1 =: M_2 :=: M_3 :=: \ldots :=: M_k \equiv N$
   de $M_0 =: M_1$ et $M_1 =: M_2$, on déduit par (tr) que $M_0 =: M_2$, donc:
   $M \equiv M_0 =: M_2 :=: M_3 :=: \ldots :=: M_k \equiv N$
5. $M \equiv M_0 =: M_1 := M_2 := M_3 :=: \ldots :=: M_k \equiv N$
   de $M_1 := M_2$ et $M_2 =: M_3$, on déduit par (tr) que $M_1 =: M_3$, donc:
   $M \equiv M_0 =: M_1 := M_3 :=: \ldots :=: M_k \equiv N$
6. $M \equiv M_0 =: M_1 := M_2 =: M_3 :=: \ldots :=: M_k \equiv N$
   de $M_0 =: M_1 := M_2$, on déduit par le lemme C-R $M_0 := Z =: M_2$, donc:
   $M \equiv M_0 := Z =: M_2 =: M_3 :=: \ldots :=: M_k \equiv N$
   de $Z =: M_2$ et $M_2 =: M_3$, on déduit par (tr) que $Z =: M_3$, donc:
   $M \equiv M_0 := Z =: M_3 :=: \ldots :=: M_k \equiv N$

- M := L := N
  on prend Z ≡ N
- M := L =: N
  on prend Z ≡ L
- M =: L =: N
  on prend Z ≡ M
- M =: L := N
  on prend Z donné par le lemme C-R

Unicité de la forme normale

- On a déjà vu avec le lemme de CHURCH-ROSSER que la forme normale d'un terme, lorsqu'elle existe, est unique.

- THEOREME. SOIT P ET Q DEUX FORMES NORMALES TELLES QUE P = Q ALORS P ET Q SONT IDENTIQUES.

- Démonstration. Si P = Q alors il existe Z tel que P := Z et Q := Z. Mais P étant une forme normale, il n'y a plus rien à réduire donc P ≡ Z. De la même manière, on a Q ≡ Z.

L'ensemble des formes normales

L'ensemble des formes normales peut être défini par induction :

- S, K et les variables sont des formes normales ; (BASE)
- Si X et Y sont des formes normales, (K X), (S X) et (S X Y) sont des formes normales ; (PAS)
- Si $X_1,\ldots,X_n$ sont des formes normales et x une variable alors x $X_1,\ldots,X_n$ est une forme normale. (PAS)

Démonstration : on commence par vérifier que les éléments de cet ensemble sont des formes normales. Puis on vérifie que les formes normales sont éléments de cet ensemble. Pour cela, on se rappelle que tout terme peut s'écrire sous la forme a $M_1\ldots M_n$ avec a atomique.
A faire en exercice.

RESOLUBILITE

- Le fait qu'un terme ne soit pas normalisable ne l'empêche pas d'avoir une valeur opératoire en tant que fonction.
- Exemple. $S\ K\ \Omega$ n'est pas normalisable puisque Ω ne l'est pas. Pourtant :

  $S\ K\ \Omega\ I := K\ I\ (\Omega\ I) := I$

- Un terme T est dit *résoluble* s'il existe $n \geq 0$ et $X_1,\ldots,X_n$ tels que $T\ X_1 \ldots X_n$ soit normalisable.

RESOLUBILITE

Théorème. Soit T un terme, il y a équivalence entre :

1. Il existe $n \geq 0$ et $X_1,\ldots,X_n$ tels que T $X_1 \ldots X_n$ soit normalisable.
2. Il existe $n \geq 0$ et $X_1,\ldots,X_n$ tels que T $X_1 \ldots X_n := I$.
3. Pour tout X, il existe $n \geq 0$ et $X_1,\ldots,X_n$ tels que T $X_1 \ldots X_n := X$.

Démonstration :

- 2 implique 3 est évident.
- 3 implique 2 est évident.
- 3 implique 1 est évident.
- 1 implique 3 se démontre en utilisant la définition inductive des formes normales.

A faire en exercice.

Le sens des termes

- Est-ce que tous les termes ont du sens ? Même s'ils n'ont pas de forme normale ?

- *Réponse de H.B. CURRY* : oui, tous les termes ont du sens. L'interprétation d'un terme est son arbre de réduction qui est une entité mathématique. C'est une approche FORMALISTE.

- *Réponse de A. CHURCH* : seul les termes ayant une forme normale ont du sens.

  Malheureusement, si l'on introduit un terme $\perp$ représentant l'indéfini et si pose l'axiome :

  $M = \perp$ si M n'a pas de forme normale

  alors on aboutit à une contradiction :

  $S\ (K\ (K\ X))\ (K\ \Omega) = S\ (K\ (K\ Y))\ (K\ \Omega)$

  donc : $S\ (K\ (K\ X))\ (K\ \Omega)\ I = S\ (K\ (K\ Y))\ (K\ \Omega)\ I$

  donc : $K\ (K\ X)\ I\ (K\ \Omega\ I) = K\ (K\ Y)\ I\ (K\ \Omega\ I)$

  donc : $K\ X\ \Omega = K\ Y\ \Omega$

  donc : $X = Y$

- Réponse de H. BARENDREGT : seuls les termes résolubles ont un sens.

  Si l'on introduit un terme $\perp$ représentant l'indéfini et si pose l'axiome :

  $M = \perp$ si M n'est pas résoluble

  alors la théorie reste COHERENTE.

L'égalité extensionnelle

- L'égalité que nous avons vu jusqu'à présent est l'égalité « faible » ou égalité entre processus de calcul.
- On a S K I x = K x (I x) = x = I x. S K I et I sont égaux en tant que fonctions mais ils ne sont pas égaux en tant que processus de calcul, c'est-à-dire que l'on n'a pas S K I = I.
- L'égalité entre fonction graphe est aussi appelée égalité extensionnelle ou égalité forte et parfois notée $X =_\eta Y$.

Les axiomes de l'égalité extensionnelle

Trois axiomes équivalents :

- *Axiome 1.* Si $F\ Z = G\ Z$ pour tout Z alors $F = G$.
- *Axiome 2.* Si $F\ x = G\ x$ et $x \notin (F\ G)$ alors $F = G$.
- *Axiome 3.* S'il existe $n \geq 0$ et des variables $x_1,\ldots,x_n \notin (F\ G)$ telles que $F\ x_1 \ldots x_n = G\ x_1 \ldots x_n$, alors $F = G$.

Exercice : démontrer que ces trois axiomes sont équivalents.

L'opérateur de point-fixe

- Soit F une fonction et X tel que F X = X, on dit que X est un *point-fixe* de F.
- On pose $Y \equiv_{def} W\ S\ (B\ W\ B)$. Y est appelé *l'opérateur de point-fixe de CURRY* ou encore *le combinateur paradoxal de CURRY*.

- THEOREME. SOIT F UN TERME, (Y F) EST UN POINT FIXE DE F. ON DIT QUE Y EST UN OPERATEUR DE POINT-FIXE.

  Etonnant, non ?

- Y F = W S (B W B) F
  = S (B W B) (B W B) F
  = B W B F (B W B F)
  = W (B F) (B W B F)
  = B F (B W B F) (B W B F)
  = F (B W B F (B W B F))
  = F (Y F)

- Autrement dit, la théorie des combinateurs nous permet de calculer le point-fixe de n'importe quelle fonction exprimable dans la théorie des combinateurs, c'est-à-dire de toutes les fonctions effectivement calculables.

L'opérateur de point-fixe

- On remarque que des fonctions calculables comme la fonction successeur n'ont pas de point-fixe….

- *Réponse classique.* On admet que nos opérateurs ne sont qu'une partie des fonctions mathématiques classiques et on cherche à les interpréter comme les fonctions d'une famille de fonctions classiques qui ont des points-fixes.

C'est ce qu'à fait D. SCOTT avec
les Modèles de SCOTT.

- *Réponse de S. FEFERMAN*. On pense au contraire que ce sont les fonctions mathématiques classiques, bridées par la théorie des ensembles, qui ne sont qu'un cas particulier de nos opérateurs et que le langages des combinateurs, fonctions processus, est un langage UNIVERSEL permettant de décrire les objets des mathématiques classiques mais aussi d'autres objets qui n'y appartiennent pas. De sorte que les mathématiques classiques peuvent ne pas percevoir que certains objets sont points- fixes de fonctions parce que ces objets sont hors de leur portée.

Les opérateurs de point-fixe

- Théorème. Tout point-fixe de (S I) est un opérateur de point-fixe.
  - Démonstration : *à faire en exercice*.

- Théorème. Si Y est un opérateur de point-fixe, Y (S I) est aussi un opérateur de point-fixe.
  - Démonstration : *à faire en exercice*.

- Conjecture de CURRY : Si $m \neq n$, $Y\ (S\ I)^n \neq Y\ (S\ I)^m$

Un dernier pour le bestiaire…

- Posons $K^{\infty} \equiv_{def} Y\ K$ alors

  $K^{\infty}\ X = Y\ K\ X = K\ (Y\ K)\ X = Y\ K = K^{\infty}$

- K^{∞} est appelé l'ABSORBEUR.