CONTENTS

INTRODUCTION

The lambda-calculus was invented in the early 1930's, by A. Church, and has been considerably developed since then. This book is an introduction to some aspects of the theory today : pure lambda-calculus, combinatory logic, semantics (models) of lambda-calculus, type systems. All these areas will be dealt with, only partially, of course, but in such a way, I think, as to illustrate their interdependence, and the essential unity of the subject.
No specific knowledge is required from the reader, but some familiarity with mathematical logic is expected ; in Chapter II, the concept of recursive function is used ; parts of Chapters VI and VII, as well as Chapter IX, involve elementary topics in predicate calculus and model theory.

For about fifteen years, the typed lambda-calculus has provoked a great deal of interest, because of its close connections with programming languages, and of the link that it establishes between the concept of program and that of intuitionistic proof : this is known as the "Curry-Howard correspondence". After the first type system, which was Curry's, many others appeared : for example, de Bruijn's Automath system, Girard's system $\mathcal{F}$, Martin-Löf's theory of intuitionistic types, Coquand-Huet's theory of constructions, Constable's Nuprl system...
This book will first introduce Coppo and Dezani's intersection type system. Here it will be called "system $\mathcal{D}\Omega$", and will be used to prove some fundamental theorems of pure lambda-calculus. It is also connected with denotational semantics : in Engeler and Scott's models, the interpretation of a term is essentially the set of its types. Next, Girard's system $\mathcal{F}$ of second order types will be considered, together with a simple extension, denoted by FA_2 (second order functional arithmetic). These types have a very transparent logical structure, and a great expressive power. They allow the Curry-Howard correspondence to be seen clearly, as well as the possibilities, and the difficulties, of using these systems as programming languages.

Chapter I

SUBSTITUTION AND BETA-CONVERSION

The terms of the λ-calculus (also called *λ-terms*) are finite sequences formed with the symbols "(,),λ" and with variables x,y,... (the set of variables is assumed to be countable). They are obtained by applying, a finite number of times, the following rules :

any variable x is a λ-term ;
whenever t and u are λ-terms, then so is (u)t ;
whenever t is a λ-term and x is a variable, then λx t is a λ-term.

The set of all terms of the λ-calculus will be denoted by L.

The term (u)t should be thought of as "u applied to t" ; it will also be denoted by ut if there is no ambiguity ; the term $(...(((u)t_1)t_2)...)t_k$ will also be written $ut_1t_2...t_k$ or $(u)t_1t_2...t_k$.
By convention, when k = 0, $(u)t_1t_2...t_k$ will denote the term u.

The *free occurrences* of a variable x in a term t are defined, by induction, as follows :
if t is the variable x, then the occurrence of x in t is free ;
if t = (u)v, then the free occurrences of x in t are those of x in u and v ;
if t = λy u, the free occurrences of x in t are those of x in u, except if x = y ; in that case, no occurrence of x in t is free.

A *free variable* in t is a variable which has at least one free occurrence in t. A term which has no free variable is called a *closed term.*
A *bound variable* in t is a variable which occurs in t just after the symbol λ.

1. Simple substitution

Let $u,t_1,\dots,t_k$ be terms and $x_1,\dots,x_k$ *distinct* variables ; the term $u\langle t_1/x_1,\dots,t_k/x_k\rangle$ is defined as the result of the replacement of every free occurrence of x_i in u by t_i $(1 \leq i \leq k)$. The definition is by induction on u, as follows :

if $u = x_i$ $(1 \leq i \leq k)$, then $u\langle t_1/x_1,\dots,t_k/x_k\rangle = t_i$;
if u is a variable and $u \neq x_1,\dots,x_k$, then $u\langle t_1/x_1,\dots,t_k/x_k\rangle = u$;
if $u = (w)v$, then $u\langle t_1/x_1,\dots,t_k/x_k\rangle = (w\langle t_1/x_1,\dots,t_k/x_k\rangle)v\langle t_1/x_1,\dots,t_k/x_k\rangle$;
if $u = \lambda x_i v$ $(1 \leq i \leq k)$, then
$u\langle t_1/x_1,\dots,t_k/x_k\rangle = \lambda x_i v\langle t_1/x_1,\dots,t_{i-1}/x_{i-1},t_{i+1}/x_{i+1},\dots,t_k/x_k\rangle$;
if $u = \lambda y\ v$, with $y \neq x_1,\dots,x_k$, then
$u\langle t_1/x_1,\dots,t_k/x_k\rangle = \lambda y\ v\langle t_1/x_1,\dots,t_k/x_k\rangle$.

Such a substitution will be called a "simple" one, in order to distinguish it from the substitution defined further on, which needs a change of bound variables. Simple substitution corresponds, in computer science, to the notion of "macro-instruction".

With the notation $u\langle t_1/x_1,\dots,t_k/x_k\rangle$, it is understood that $x_1,\dots,x_k$ are distinct variables. Moreover, their order does not matter ; in other words :

$u\langle t_1/x_1,\dots,t_k/x_k\rangle = u\langle t_{\sigma 1}/x_{\sigma 1},\dots,t_{\sigma k}/x_{\sigma k}\rangle$ for any permutation σ of $\{1,\dots,k\}$.
The proof is immediate by induction on the length of u ; also immediate is the following :

If each of the terms $t_1,\dots,t_k$ are variables, then the term $u\langle t_1/x_1,\dots,t_k/x_k\rangle$ has the same length as u.

Lemma 1. **Let $v,t_1,\dots,t_k,u_1,\dots,u_l$ be λ-terms, and $x_1,\dots,x_k,y_1,\dots,y_l$ distinct variables. If $y_1,\dots,y_l$ are not free in v, then :**
$v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle = v\langle t_1/x_1,\dots,t_k/x_k\rangle$.

Proof by induction on v. The result is clear when v is either a variable or a term of the form $(v_1)v_2$. Now suppose $v = \lambda x\ w$; then :

if $x = x_i$, say x_1, then :
$v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle = \lambda x_1 w\langle t_2/x_2,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle =$
$\lambda x_1 w\langle t_2/x_2,\dots,t_k/x_k\rangle$ (by induction hypothesis) $= v\langle t_1/x_1,\dots,t_k/x_k\rangle$;
if $x = y_j$, say y_1, then :
$v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle = \lambda y_1 w\langle t_1/x_1,\dots,t_k/x_k,u_2/y_2,\dots,u_l/y_l\rangle$
$= \lambda y_1 w\langle t_1/x_1,\dots,t_k/x_k\rangle$ by induction hypothesis, (since $y_2,\dots,y_l$ are not free in w) $= v\langle t_1/x_1,\dots,t_k/x_k\rangle$.

if $x \neq x_1,\dots,x_k,y_1,\dots,y_l$, then :
$v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle = \lambda x\ w\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$
$= \lambda x\ w\langle t_1/x_1,\dots,t_k/x_k\rangle$ (by induction hypothesis) $= v\langle t_1/x_1,\dots,t_k/x_k\rangle$.
Q.E.D.

Lemma 2. **Let $v,t_1,\dots,t_k,u_1,\dots,u_l$ be λ-terms, and $x_1,\dots,x_k,y_1,\dots,y_l$ distinct variables. If $y_1,\dots,y_l$ are not free in $t_1,\dots,t_k$, then :**
$v\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle = v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$.

Proof by induction on the length of v :
i) v is a variable :
if $v = x_i$, then the identity to be proved is $t_i\langle u_1/y_1,\dots,u_l/y_l\rangle = t_i$, which follows from lemma 1, since $y_1,\dots,y_l$ are not free in t_i ;
if $v = y_j$, or $v \neq x_1,\dots,x_k,y_1,\dots,y_l$, then the result is clear.
ii) $v = (v_1)v_2$; the result is obvious, by applying the induction hypothesis to v_1,v_2.
iii) $v = \lambda x\ w$;
if $x = x_i$, say x_1, then :
$v\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle = \lambda x_1 w\langle t_2/x_2,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle$
$= \lambda x_1 w\langle t_2/x_2,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$ by induction hypothesis,
$= v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$;
if $x = y_j$, say y_1, then :
$v\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle = \lambda y_1 w\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_2/y_2,\dots,u_l/y_l\rangle$
$= \lambda y_1 w\langle t_1/x_1,\dots,t_k/x_k,u_2/y_2,\dots,u_l/y_l\rangle$ by induction hypothesis,
$= v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$;
if $x \neq x_1,\dots,x_k,y_1,\dots,y_l$, then :
$v\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle = \lambda x\ w\langle t_1/x_1,\dots,t_k/x_k\rangle\langle u_1/y_1,\dots,u_l/y_l\rangle$
$= \lambda x\ w\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$ by induction hypothesis,
$= v\langle t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l\rangle$.
Q.E.D.

Lemma 3. **Let $u,t_1,\dots,t_k$ be λ-terms, and $\{x_1,\dots,x_k\}$, $\{y_1,\dots y_k\}$ two sets of variables such that none of the y_i's occur in u. Then :**
$u\langle y_1/x_1,\dots,y_k/x_k\rangle\langle t_1/y_1,\dots,t_k/y_k\rangle = u\langle t_1/x_1,\dots,t_k/x_k\rangle$.

Proof by induction on u. If u is a variable (and thus $u \neq y_1,\dots,y_k$), then the conclusion is obvious. So is it in the case $u = (w)v$. If $u = \lambda z\ v$, then $z \notin \{y_1,\dots,y_k\}$. If z is an x_i, say x_1, then $u\langle y_1/x_1,\dots,y_k/x_k\rangle\langle t_1/y_1,\dots,t_k/y_k\rangle = (\lambda x_1 v\langle y_2/x_2,\dots,y_k/x_k\rangle)\langle t_1/y_1,\dots,t_k/y_k\rangle =$ $\lambda x_1 v\langle y_2/x_2,\dots,y_k/x_k\rangle\langle t_2/y_2,\dots,t_k/y_k\rangle$ (because of the assumption that y_1 does not occur in $\lambda x_1 v$, and therefore, neither in $\lambda x_1 v\langle y_2/x_2,\dots,y_k/x_k\rangle$) $= \lambda x_1 v\langle t_2/x_2,\dots,t_k/x_k\rangle$ (by

induction hypothesis) $= u\langle t_1/x_1,\ldots,t_k/x_k\rangle$.
If $z \notin \{x_1,\ldots,x_k\}$, then :
$u\langle y_1/x_1,\ldots,y_k/x_k\rangle\langle t_1/y_1,\ldots,t_k/y_k\rangle = \lambda z\, v\langle y_1/x_1,\ldots,y_k/x_k\rangle\langle t_1/y_1,\ldots,t_k/y_k\rangle$
$= \lambda z\, v\langle t_1/x_1,\ldots,t_k/x_k\rangle$ (induction hypothesis) $= u\langle t_1/x_1,\ldots,t_k/x_k\rangle$.
Q.E.D.

Let R be a binary relation on L ; we will say that R is *λ-compatible* if it is reflexive and satisfies :

$$t\, R\, t' \Rightarrow \lambda x\, t\, R\, \lambda x\, t' \; ; \; t\, R\, t',\, u\, R\, u' \Rightarrow (u)t\, R\, (u')t' .$$

Lemma 4. If R is λ-compatible and $t_1\, R\, t'_1,\ldots,t_k\, R\, t'_k$, then : $u\langle t_1/x_1,\ldots,t_k/x_k\rangle\, R\, u\langle t'_1/x_1,\ldots,t'_k/x_k\rangle$.

Immediate proof by induction on the length of u.
Q.E.D.

**Proposition 1. Let R be a binary relation on L. Then, the least λ-compatible binary relation ρ containing R is defined by the following condition :
(1) $t\, \rho\, t' \Leftrightarrow$ there exists terms $T, t_1,\ldots,t_k, t'_1,\ldots,t'_k$ and distinct variables $x_1,\ldots,x_k$ such that : $t_i\, R\, t'_i$ $(1 \le i \le k)$ and $t = T\langle t_1/x_1,\ldots,t_k/x_k\rangle$, $t' = T\langle t'_1/x_1,\ldots,t'_k/x_k\rangle$.**

Let ρ' be the least λ-compatible binary relation containing R, and ρ the relation defined by condition (1) above. It follows from the previous lemma that $\rho' \supset \rho$. It is easy to see that $\rho \supset R$ (take $T = x_1$). It thus remains to prove that ρ is λ-compatible.
By taking $k = 0$ in condition (1), we see that ρ is reflexive.
Suppose $t = T\langle t_1/x_1,\ldots,t_k/x_k\rangle$, $t' = T\langle t'_1/x_1,\ldots,t'_k/x_k\rangle$. Let $y_1,\ldots,y_k$ be distinct variables not occurring in T. Let $V = T\langle y_1/x_1,\ldots,y_k/x_k\rangle$. Then, it follows from lemma 3 that $t = V\langle t_1/y_1,\ldots,t_k/y_k\rangle$ and $t' = V\langle t'_1/y_1,\ldots,t'_k/y_k\rangle$. Thus the distinct variables $x_1,\ldots,x_k$ in condition (1) can be arbitrarily chosen, except in some finite set.
Now suppose $t\, \rho\, t'$ and $u\, \rho\, u'$; then :
$t = T\langle t_1/x_1,\ldots,t_k/x_k\rangle$, $t' = T\langle t'_1/x_1,\ldots,t'_k/x_k\rangle$ with $t_i\, R\, t'_i$;
$u = U\langle u_1/y_1,\ldots,u_l/y_l\rangle$, $u' = U\langle u'_1/y_1,\ldots,u'_l/y_l\rangle$ with $u_j\, R\, u'_j$.
By the previous remark, we can assume that $x_1,\ldots,x_k,y_1,\ldots,y_l$ are distinct, different from x, and also that none of the x_i's occur in U, and none of the y_j occur in T. Therefore :
$\lambda x\, t = (\lambda x\, T)\langle t_1/x_1,\ldots,t_k/x_k\rangle$, $\lambda x\, t' = (\lambda x\, T)\langle t'_1/x_1,\ldots,t'_k/x_k\rangle$, which proves that $\lambda x\, t\, \rho\, \lambda x\, t'$.
Also, by lemma 1 : $t = T\langle t_1/x_1,\ldots,t_k/x_k,u_1/y_1,\ldots,u_l/y_l\rangle$,
$t' = T\langle t'_1/x_1,\ldots,t'_k/x_k,u'_1/y_1,\ldots,u'_l/y_l\rangle$ (since none of the y_j's occur in T) ;
and similarly $u = U\langle t_1/x_1,\ldots,t_k/x_k,u_1/y_1,\ldots,u_l/y_l\rangle$,

$u' = U\langle t'_1/x_1,\ldots,t'_k/x_k,u'_1/y_1,\ldots,u'_l/y_l\rangle$ (since none of the x_i's occur in U).
Let $V = (U)T$; then $(u)t = V\langle t_1/x_1,\ldots,t_k/x_k,u_1/y_1,\ldots,u_l/y_l\rangle$,
$(u')t' = V\langle t'_1/x_1,\ldots,t'_k/x_k,u'_1/y_1,\ldots,u'_l/y_l\rangle$ and thus $(u)t\, \rho\, (u')t'$.
Q.E.D.

2. Alpha-equivalence and substitution

We will now define an equivalence relation on the set L of all λ-terms. It is called α-equivalence, and denoted by ≡.
Intuitively, $u \equiv u'$ means that u' is obtained from u by renaming the bound variables in u ; more precisely, $u \equiv u'$ if and only if u and u' have the same sequence of symbols (when all variables are considered equal), the same free occurrences of the same variables, and if each λ binds the same occurrences of variables in u and in u'.

We define $u \equiv u'$, on L, by induction on the length of u, by the following clauses :

if u is a variable, then $u \equiv u'$ if and only if $u = u'$;
if $u = (w)v$, then $u \equiv u'$ if and only if $u' = (w')v'$, with $v \equiv v'$ and $w \equiv w'$;
if $u = \lambda x\, v$, then $u \equiv u'$ if and only if $u' = \lambda x' v'$, with $v\langle y/x\rangle \equiv v'\langle y/x'\rangle$ for all variables y except a finite number.

(Note that $v\langle y/x\rangle$ has the same length as v, thus is shorter than u, which guarantees the correctness of the inductive definition).

It can be seen immediately, by induction on the length of u, that
if $u \equiv u'$, then u and u' have the same free variables and the same length.

The relation ≡ is an equivalence relation on L.

Indeed, the proof of the three following properties, by induction on u, is trivial :
$u \equiv u$; $u \equiv u' \Rightarrow u' \equiv u$; $u \equiv u',\, u' \equiv u'' \Rightarrow u \equiv u''$.

**Proposition 2. Let $u, u', t_1,\ldots,t_k$ be λ-terms, and $x_1,\ldots,x_k$ distinct variables. If $u \equiv u'$ and if no free variable in $t_1,\ldots,t_k$ is bound in u,u', then :
$u\langle t_1/x_1,\ldots,t_k/x_k\rangle \equiv u'\langle t_1/x_1,\ldots,t_k/x_k\rangle$.**

Note that, since $u \equiv u'$, u and u' have the same free variables. Thus it can be assumed that $x_1,\ldots,x_k$ are free in u and u' : indeed, if $x_1,\ldots,x_l$ are those x_i variables which are free in u

and u', then, by lemma 1 :
$u<t_1/x_1,\dots,t_k/x_k> = u<t_1/x_1,\dots,t_l/x_l>$ and $u'<t_1/x_1,\dots,t_k/x_k> = u'<t_1/x_1,\dots,t_l/x_l>$.
The proof of the proposition proceeds by induction on u. The result is immediate if u is a variable, or $u = (w)v$. Suppose $u = \lambda x\, v$. Then $u' = \lambda x'\, v'$ and $v<y/x> \equiv v'<y/x'>$ for all variables y except a finite number.
Since $x_1,\dots,x_k$ are free in u and u', x and x' are different from $x_1,\dots,x_k$. Thus $u<t_1/x_1,\dots,t_k/x_k> = \lambda x\, v<t_1/x_1,\dots,t_k/x_k>$ and
$u'<t_1/x_1,\dots,t_k/x_k> = \lambda x'\, v'<t_1/x_1,\dots,t_k/x_k>$. Hence it is sufficient to show that $v<t_1/x_1,\dots,t_k/x_k><y/x> \equiv v'<t_1/x_1,\dots,t_k/x_k><y/x'>$ for all variables y except a finite number. Since x,x' are not free in $t_1,\dots,t_k$, it follows from lemma 2 that $v<t_1/x_1,\dots,t_k/x_k><y/x> = v<y/x,t_1/x_1,\dots,t_k/x_k>$ and
$v'<t_1/x_1,\dots,t_k/x_k><y/x'> = v'<y/x',t_1/x_1,\dots,t_k/x_k>$. If $y \neq x_1,\dots,x_k$, then, again by lemma 2 : $v<t_1/x_1,\dots,t_k/x_k><y/x> = v<y/x><t_1/x_1,\dots,t_k/x_k>$ and
$v'<t_1/x_1,\dots,t_k/x_k><y/x'> = v'<y/x'><t_1/x_1,\dots,t_k/x_k>$.
Since $v<y/x> \equiv v'<y/x'>$ for all variables y except a finite number, and $v<y/x>$ is shorter than u, the induction hypothesis gives :
$v<y/x><t_1/x_1,\dots,t_k/x_k> \equiv v'<y/x'><t_1/x_1,\dots,t_k/x_k>$, thus :
$v<t_1/x_1,\dots,t_k/x_k><y/x> \equiv v'<t_1/x_1,\dots,t_k/x_k><y/x'>$ for all variables y except a finite number.
Q.E.D.

Corollary. The relation $\equiv$ is λ-compatible.

Suppose $u \equiv u'$. We need to prove that $\lambda x\, u \equiv \lambda x\, u'$, that is to say $u<y/x> \equiv u'<y/x>$ for all variables y except a finite number. But this follows from the previous proposition, provided that y is not a bound variable in u or in u'.
Q.E.D.

Corollary. If $t_1,\dots,t_k,t'_1,\dots,t'_k$ are terms, and $x_1,\dots,x_k$ are distinct variables, then :
$t_1 \equiv t'_1,\dots,t_k \equiv t'_k \Rightarrow u<t_1/x_1,\dots,t_k/x_k> \equiv u<t'_1/x_1,\dots,t'_k/x_k>$.

This follows from the previous corollary and lemma 4.

However, note that it is not true that $u \equiv u' \Rightarrow u<t/x> \equiv u'<t/x>$. For example, $\lambda y\, x \equiv \lambda z\, x$, while $\lambda y\, y \not\equiv \lambda z\, y$. This is the reason why the simple substitution is not the appropriate one.

Lemma 5. $\lambda x\, u \equiv \lambda y\, u<y/x>$ whenever y is a variable which does not occur in u.

By lemma 3, $u<z/x> = u<y/x><z/y>$ for any variable z, since y does not occur in u. Hence the result follows from the definition of $\equiv$.
Q.E.D.

Lemma 6. Let t be a term, and $x_1,\dots,x_k$ be variables. Then there exists a term t', $t' \equiv t$, such that none of $x_1,\dots,x_k$ are bound in t'.

The proof is by induction on t. The result is immediate if t is a variable, or if $t = (v)u$. If $t = \lambda x\, u$, then, by induction hypothesis, there exists a term u', $u' \equiv u$, in which none of $x_1,\dots,x_k$ are bound. By the previous lemma, $t \equiv \lambda x\, u' \equiv \lambda y\, u'<y/x>$ with $y \neq x_1,\dots,x_k$. Thus it is sufficient to take $t' = \lambda y\, u'<y/x>$.
Q.E.D.

From now on, α-equivalent terms will be identified ; hence we will deal with the quotient set $L/\equiv$; it is denoted by Λ.
For each variable x, its equivalence class will still be denoted by x (it is actually $\{x\}$). Furthermore, the operations $u,v \to (v)u$ and $u,x \to \lambda x\, u$ are compatible with $\equiv$ and are therefore defined in Λ.
Moreover, if $u \equiv u'$, then u and u' have the same free variables. Hence it is possible to define the free variables of a member of Λ.

Consider terms $u,t_1,\dots,t_k$ and distinct variables $x_1,\dots,x_k$. The term $u[t_1/x_1,\dots,t_k/x_k]$ (being the result of the replacement of every free occurrence of x_i in u by t_i, for $i = 1,\dots,k$) is defined as $u'<t_1/x_1,\dots,t_k/x_k>$, where u' is a term such that $u' \equiv u$ and no bound variable in u' is free in $t_1,\dots,t_k$: this is possible because, first of all such a term u' exists (this has just been proved) ; then, by proposition 2, the equivalence class of $u'<t_1/x_1,\dots,t_k/x_k>$ does not depend on the choice of u'. Hence $u \equiv u' \Rightarrow u[t_1/x_1,\dots,t_k/x_k] \equiv u'[t_1/x_1,\dots,t_k/x_k]$.
Finally, it follows from a corollary of proposition 2 that :

$$t_1 \equiv t'_1,\dots,t_k \equiv t'_k \Rightarrow u[t_1/x_1,\dots,t_k/x_k] \equiv u[t'_1/x_1,\dots,t'_k/x_k].$$

So the substitution operation $u,t_1,\dots,t_k \to u[t_1/x_1,\dots,t_k/x_k]$ is well defined in Λ. It corresponds to the replacement of the free occurrences of x_i in u by t_i ($1 \leq i \leq k$), provided that a representative of u has been chosen which has no bound variable in $t_1,\dots,t_k$.
Thus the substitution operation satisfies the following lemmas, already stated for the simple substitution :

Lemma 7. Let $v,t_1,\dots,t_k,u_1,\dots,u_l$ be λ-terms, and $x_1,\dots,x_k,y_1,\dots,y_l$ distinct variables. If $y_1,\dots,y_l$ are not free in v, then :
$v[t_1/x_1,\dots,t_k/x_k,u_1/y_1,\dots,u_l/y_l] \equiv v[t_1/x_1,\dots,t_k/x_k]$.

Lemma 8. Let $v,t_1,\ldots,t_k,u_1,\ldots,u_l$ be λ-terms, and $x_1,\ldots,x_k,y_1,\ldots,y_l$ distinct variables. If $y_1,\ldots,y_l$ are not free in $t_1,\ldots,t_k$, then :
$v[t_1/x_1,\ldots,t_k/x_k][u_1/y_1,\ldots,u_l/y_l] \equiv v[t_1/x_1,\ldots,t_k/x_k,u_1/y_1,\ldots,u_l/y_l]$.

Lemma 9. Let u, $t_1,\ldots,t_k$ be terms, and $\{x_1,\ldots,x_k\}$, $\{y_1,\ldots y_k\}$ two sets of variables, such that none of the y_i's are free in u. Then :

$$u[y_1/x_1,\ldots,y_k/x_k][t_1/y_1,\ldots,t_k/y_k] \equiv u[t_1/x_1,\ldots,t_k/x_k].$$

However, the following lemma would not be true for the simple substitution.

Lemma 10. Let $u,a_1,\ldots,a_k,b_1,\ldots,b_l$ be terms, and $x_1,\ldots,x_k,y_1,\ldots,y_l$ distinct variables. Then :

$$u[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k] \equiv u[a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l],$$

where $b'_j = b_j[a_1/x_1,\ldots,a_k/x_k]$ $(1 \le j \le l)$.
Moreover, if $y_1,\ldots,y_l$ are not free in $a_1,\ldots,a_k$, then :

$$u[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k] \equiv u[a_1/x_1,\ldots,a_k/x_k][b'_1/y_1,\ldots,b'_l/y_l].$$

The second part of the lemma follows from the first one and from lemma 8.
The proof of the first part proceeds by induction on u. The result is immediate if u is a variable (examine the cases where that variable is x_i or y_j), or if $u = (w)v$.
Suppose $u = \lambda z\, v$. Choose representatives of $u,a_1,\ldots,a_k,b_1,\ldots,b_l$ with bound variables different from $x_1,\ldots,x_k,y_1,\ldots,y_l$, and not free in $u,a_1,\ldots,a_k,b_1,\ldots,b_l$. Then :
$u[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k] \equiv \lambda z\, v\langle b_1/y_1,\ldots,b_l/y_l\rangle\langle a_1/x_1,\ldots,a_k/x_k\rangle$
$\equiv \lambda z\, v[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k]$.
By induction hypothesis,
$v[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k] \equiv v[a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l]$, which is equivalent to $v\langle a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l\rangle$, according to the choice of the representatives of v,a_i,b_j.
Indeed, $b'_j \equiv b_j[a_1/x_1,\ldots,a_k/x_k] \equiv b_j\langle a_1/x_1,\ldots,a_k/x_k\rangle$, hence the free variables in b'_j are not bound in v.
Therefore :
$u[b_1/y_1,\ldots,b_l/y_l][a_1/x_1,\ldots,a_k/x_k] \equiv \lambda z\, v\langle a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l\rangle \equiv$
$(\lambda z\, v)\langle a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l\rangle \equiv u[a_1/x_1,\ldots,a_k/x_k,b'_1/y_1,\ldots,b'_l/y_l]$.
Q.E.D.

As indicated above, this result does not hold for the simple substitution :
if $u = \lambda x\, y$, $b = x$, then $u\langle a/x,b'/y\rangle = \lambda x\, b' = \lambda x\, a$ (since $b' = b[a/x]$), while $u\langle b/y\rangle\langle a/x\rangle = \lambda x\, x$.

3. Beta-conversion

Let R be a binary relation (on an arbitrary set) ; the least transitive binary relation which contains R is obviously the relation R' defined by :
t R' u $\Leftrightarrow$ there exist an integer n and terms $v_0 = t,v_1,\ldots,v_{n-1},v_n = u$ such that v_i R v_{i+1} $(0 \le i < n)$.
R' is called the *transitive closure* of R.
The *Church-Rosser* (C.-R.) *property* is satisfied by a relation R if and only if :
t R u , t R u' $\Rightarrow$ for some v, u R v and u' R v.

Lemma. Let R be a binary relation which satisfies the Church-Rosser property. Then the transitive closure of R also satisfies it.

Let R' be that transitive closure. We will first prove the following property :
t R' u , t R u' $\Rightarrow$ for some v, u R v and u' R' v .
t R' u means that there exists a sequence $v_0 = t,v_1,\ldots,v_{n-1},v_n = u$ such that v_i R v_{i+1} $(0 \le i < n)$.
The proof is by induction on n ; the case n = 1 is just the hypothesis of the lemma.
Now since t R' v_{n-1} and t R u', for some w, v_{n-1} R w and u' R' w. But v_{n-1} R u, so u R v and w R v for some v (C.-R. property for R). Therefore u' R' v, which gives the result.
Now we can prove the lemma : the assumption is t R' u and t R' u', so there exists a sequence : $(v_0 = t),v_1,\ldots,v_{n-1},(v_n = u')$ such that v_i R v_{i+1} $(0 \le i < n)$.
The proof is by induction on n : the case n = 1 has just been settled.
Since t R' u and t R' v_{n-1}, by induction hypothesis, we have u R' w and v_{n-1} R' w for some w. Now v_{n-1} R u', so, by the previous property, w R v and u' R' v for some v. Thus u R' v.
Q.E.D.

Now we will consider binary relations on Λ. Such relations are identified with binary relations on L which are compatible with the equivalence relation ≡ .

Proposition 3. If $(\lambda x\, u)t \equiv (\lambda x'\, u')t'$, then $u[t/x] \equiv u'[t'/x']$.

By definition of ≡, the assumption is : $t \equiv t'$ and $\lambda x\, u \equiv \lambda x'\, u'$. Thus $u\langle y/x\rangle \equiv u'\langle y/x'\rangle$ for all variables y except a finite number. Choose representatives of u,u' with no bound variable free in t or t'. Then, by proposition 2, $(u\langle y/x\rangle)\langle t/y\rangle \equiv (u'\langle y/x'\rangle)\langle t/y\rangle$. Let y be a variable which does not occur in u,u'. It then follows from lemma 3 that $u\langle t/x\rangle \equiv u'\langle t/x'\rangle$; moreover $u'\langle t/x'\rangle \equiv u'\langle t'/x'\rangle$ (corollary of proposition 2). Thus $u\langle t/x\rangle \equiv u'\langle t'/x'\rangle$, that is $u[t/x] \equiv u'[t'/x']$.
Q.E.D.

A term of the form $(\lambda x\ u)t$ is called a ***redex***, $u[t/x]$ is called its ***contractum***. The previous proposition shows that this notion is defined on Λ.
A binary relation β_0 will now be defined on Λ ; $t\ \beta_0\ t'$ should be read as : "t' is obtained by contracting a redex (or by a ***β-reduction***) in t ". The definition is by induction on t :

if t is a variable, then there is no t' such that $t\ \beta_0\ t'$;
if $t = \lambda x\ u$, then $t\ \beta_0\ t'$ if and only if $t' = \lambda x\ u'$, with $u\ \beta_0\ u'$;
if $t = (v)u$, then $t\ \beta_0\ t'$ if and only if :
either $t' = (v)u'$ with $u\ \beta_0\ u'$,
or $t' = (v')u$ with $v\ \beta_0\ v'$,
or else $v = \lambda x\ w$ and $t' = w[u/x]$.

It is clear from this definition that, whenever $t\ \beta_0\ t'$, any free variable in t' is also free in t.

The ***β-conversion*** is the least binary relation β on Λ, which is reflexive, transitive, and contains β_0 . So $t\ \beta\ t'$ if and only if there exists a sequence $(t_0 = t),t_1,..,t_{n-1},(t_n = t')$ such that $t_i\ \beta_0\ t_{i+1}$ for $1 \leq i \leq n-1$ $(n \geq 0)$.

Therefore, whenever $t\ \beta\ t'$, any free variable in t' is also free in t.
The next two propositions give two simple characterizations of β.

Proposition 4. The β-conversion is the least transitive λ-compatible binary relation β such that $(\lambda x\ u)t\ \beta\ u[t/x]$ for all terms t,u and variable x.

Clearly, $t\ \beta_0\ t'$, $u\ \beta_0\ u' \Rightarrow \lambda x\ t\ \beta_0\ \lambda x\ t'$ and $(u)t\ \beta\ (u')t'$. Hence β is λ-compatible. Conversely, if R is a λ-compatible binary relation and if $(\lambda x\ u)t\ R\ u[t/x]$ for all terms t,u, then it follows immediately from the definition of β_0 that $R \supset \beta_0$ (prove $t\ \beta_0 t' \Rightarrow t\ R\ t'$ by induction on t). So, if R is transitive, then $R \supset \beta$.
Q.E.D.

Proposition 5. β is the transitive closure of the binary relation ρ defined on L by : $u\ \rho\ u' \Leftrightarrow$ there exist a term v and redexes $t_1,...,t_k$ with contractums $t'_1,...,t'_k$ such that $u \equiv v<t_1/x_1,...,t_k/x_k>$, $u' \equiv v<t'_1/x_1,...,t'_k/x_k>$.

Indeed, according to proposition 1, ρ is the least λ-compatible binary relation on L which is compatible with $\equiv$, and such that $t\ \rho\ t'$ for any redex t with contractum t'. Thus, by the previous proposition, β is the transitive closure of ρ.
Q.E.D.

Proposition 6. If $t\ \beta\ t'$ and $u\ \beta\ u'$ then $u[t/x]\ \beta\ u'[t'/x]$.

If $u = u'$, then the result is given by lemma 4, since the relation β is λ-compatible (it is assumed that the chosen representatives of u and u' have no bound variable free in t or t'). Hence it is sufficient to show that : $t\ \beta\ t'$ and $u\ \beta_0\ u' \Rightarrow u[t/x]\ \beta\ u'[t'/x]$.
This is done by induction on the length of u. It follows from the definition of β_0 that the different possibilities for u,u' are :
i) $u = \lambda y\ v$, $u' = \lambda y\ v'$, and $v\ \beta_0\ v'$. Then, by induction hypothesis, $v[t/x]\ \beta\ v'[t'/x]$, and, since β is λ-compatible, $\lambda y\ v[t/x]\ \beta\ \lambda y\ v'[t'/x]$, that is $u[t/x]\ \beta\ u'[t'/x]$.
ii) $u = (w)v$ and $u' = (w)v'$, with $v\ \beta_0\ v'$. Then, by induction hypothesis, $v[t/x]\ \beta\ v'[t'/x]$; and, by lemma 4, $w[t/x]\ \beta\ w[t'/x]$. Therefore $(w[t/x])v[t/x]\ \beta\ (w[t'/x])v'[t'/x]$, since β is λ-compatible ; that is to say $u[t/x]\ \beta\ u'[t'/x]$.
iii) $u = (w)v$ and $u' = (w')v$, with $w\ \beta_0\ w'$. Same proof.
iv) $u = (\lambda y\ w)v$ and $u' = w[v/y]$. Then $u[t/x] = (\lambda y\ w[t/x])v[t/x]$ (a representative of u has been chosen in which neither x, nor any free variable of t are bound). Thus $u[t/x]\ \beta_0\ w[t/x][v[t/x]/y]$; now, by lemma 10 : $w[t/x][v[t/x]/y] \equiv w[v/y][t/x] \equiv u'[t/x]$, and by lemma 4, $u'[t/x]\ \beta\ u'[t'/x]$. Hence $u[t/x]\ \beta\ u'[t'/x]$.
Q.E.D.
This proposition is also an immediate consequence of lemma 12.

A term t is said to be ***normal***, or to be ***in normal form***, if it contains no redex.
So the normal terms are those which are obtained by applying, a finite number of times, the following rules :

any variable x is a normal term ;
whenever t is normal, so is $\lambda x\ t$;
if t,u are normal and if the first symbol in u is not λ, then $(u)t$ is normal.

This definition yields, immediately, the following properties :
A term is normal if and only if it is of the form $\lambda x_1...\lambda x_k(x)t_1...t_n$ $(k,n \geq 0)$, where x is a variable and $t_1,...,t_n$ are normal terms.
A term t is normal if and only if there is no term t' such that $t\ \beta_0\ t'$.

Thus a normal term is "minimal" with respect to β, which means that, whenever t is normal, $t\ \beta\ t' \Rightarrow t \equiv t'$. However the converse is not true : take $t = (\lambda x(x)x)\lambda x(x)x$, then $t\ \beta\ t' \Rightarrow t \equiv t'$ although t is not normal.

A term t is said to be ***normalizable*** if $t\ \beta\ t'$ for some normal term t'. A term t is said to be ***strongly normalizable*** if there is no infinite sequence $t_0 = t,t_1,...,t_n$,... such that $t_i\ \beta_0\ t_{i+1}$ for all $i \geq 0$ (the term t is then obviously normalizable).
For example, $\lambda x\ x$ is a normal term, $(\lambda x(x)x)\lambda x\ x$ is strongly normalizable, $\omega = (\lambda x(x)x)\lambda x(x)x$ is not normalizable, $(\lambda x\ y)\omega$ is normalizable but not strongly normalizable.

For normalizable terms, the problem of the uniqueness of the normal form arises. It is solved by the following theorem :

Church-Rosser theorem. The β-conversion satisfies the Church-Rosser property.

This yields the uniqueness of the normal form : if $t\ \beta\ t_1$, $t\ \beta\ t_2$, with t_1,t_2 normal, then, according to the theorem, there exists a term t_3 such that $t_1\ \beta\ t_3$, $t_2\ \beta\ t_3$. Thus $t_1 \equiv t_3 \equiv t_2$.

In order to prove that β satisfies the Church-Rosser property, it is sufficient to exhibit a binary relation ρ on Λ which satisfies the Church-Rosser property and has the β-conversion as its transitive closure.
One could think of taking ρ to be the "reflexive closure" of β_0, which would be defined by $x\ \rho\ y \Leftrightarrow x = y$ or $x\ \beta_0\ y$. But this relation ρ does not satisfy the Church-Rosser property : for example, if $t = (\lambda x(x)x)r$, where r is a redex with contractum r', $u = (r)r$ and $v = (\lambda x(x)x)r'$, then $t\ \beta_0\ u$ and $t\ \beta_0\ v$, while there is no term w such that $u\ \beta_0\ w$ and $v\ \beta_0\ w$.

A suitable definition of ρ is as the least λ-compatible binary relation on Λ such that :
$t\ \rho\ t'$, $u\ \rho\ u' \Rightarrow (\lambda x\ u)t\ \rho\ u'[t'/x]$.

To prove that $\beta \supset \rho$, it is enough to see that $t\ \beta\ t'$, $u\ \beta\ u' \Rightarrow (\lambda x\ u)t\ \beta\ u'[t'/x]$; now : $(\lambda x\ u)t\ \beta\ (\lambda x\ u')t'$ (since β is λ-compatible) and : $(\lambda x\ u')t'\ \beta\ u'[t'/x]$; then the expected result follows, by transitivity.
Therefore, β contains the transitive closure ρ' of ρ. But of course $\rho \supset \beta_0$, so $\rho' \supset \beta$.
Hence β is the transitive closure of ρ. It thus remains to prove that ρ satisfies the Church-Rosser property.

By definition, ρ is the set of all pairs of (equivalence classes of) terms obtained by applying, a finite number of times, the following rules :

1) $t\ \rho\ t$;
2) $t\ \rho\ t' \Rightarrow \lambda x\ t\ \rho\ \lambda x\ t'$;
3) $t\ \rho\ t'$ and $u\ \rho\ u' \Rightarrow (u)t\ \rho\ (u')t'$;
4) $t\ \rho\ t'$, $u\ \rho\ u' \Rightarrow (\lambda x\ u)t\ \rho\ u'[t'/x]$.

Lemma 11. i) If $x\ \rho\ t'$, where x is a variable, then $t' = x$.
ii) If $\lambda x\ u\ \rho\ t'$, then $t' \equiv \lambda x\ u'$, and $u\ \rho\ u'$.
iii) If $(v)u\ \rho\ t'$, then either $t' \equiv (v')u'$ with $u\ \rho\ u'$ and $v\ \rho\ v'$, or $v \equiv \lambda x\ w$ and $t' \equiv w'[u'/x]$ with $u\ \rho\ u'$ and $w\ \rho\ w'$.

i) $x\ \rho\ t'$ could only be obtained by applying rule 1, hence $t' = x$.
ii) Consider the last rule applied to obtain $\lambda x\ u\ \rho\ t'$; the form of the term on the left shows that it is necessarily rule 1 or rule 2 ; the result then follows.
iii) Same method : the last rule applied to obtain $(v)u\ \rho\ t'$ is 1,3, or 4 ; this yields the conclusion.
Q.E.D.

Lemma 12. Whenever $t\ \rho\ t'$ and $u\ \rho\ u'$, then $u[t/x]\ \rho\ u'[t'/x]$.

The proof proceeds by induction on the length of the derivation of $u\ \rho\ u'$ by means of rules 1,2,3,4 ; consider the last rule used :

if it is rule 1, then $u \equiv u'$, and the result follows by lemma 4 ;

if it is rule 2, then $u \equiv \lambda y\ v$, $u' \equiv \lambda y\ v'$ and $v\ \rho\ v'$. It can be assumed, with a suitable choice of the representatives of u,u', that $y \neq x$. Since $t\ \rho\ t'$, the induction hypothesis implies $v[t/x]\ \rho\ v'[t'/x]$; hence $\lambda y\ v[t/x]\ \rho\ \lambda y\ v'[t'/x]$ (rule 2), that is to say $u[t/x]\ \rho\ u'[t'/x]$;

if it is rule 3, then $u \equiv (w)v$ and $u' \equiv (w')v'$, with $v\ \rho\ v'$ and $w\ \rho\ w'$. Thus, by induction hypothesis, $v[t/x]\ \rho\ v'[t'/x]$ and $w[t/x]\ \rho\ w'[t'/x]$; Therefore, by applying rule 3, we obtain $(w[t/x])v[t/x]\ \rho\ (w'[t'/x])v'[t'/x]$, that is $u[t/x]\ \rho\ u'[t'/x]$;

if it is rule 4, then $u \equiv (\lambda y\ w)v$ and $u' \equiv w'[v'/y]$, with $v\ \rho\ v'$ and $w\ \rho\ w'$. Thus, by induction hypothesis, $v[t/x]\ \rho\ v'[t'/x]$ and $w[t/x]\ \rho\ w'[t'/x]$. By rule 4, it follows that $(\lambda y\ w[t/x])v[t/x]\ \rho\ w'[t'/x][v'[t'/x]/y]$; now $\lambda y\ w[t/x] \equiv (\lambda y\ w)[t/x]$ (it can be assumed that $y \neq x$, with a suitable choice of the representative of u). Therefore, $(\lambda y\ w[t/x])v[t/x] \equiv u[t/x]$.
On the other hand, $w'[t'/x][v'[t'/x]/y] \equiv w'[v'/y][t'/x]$ (lemma 10 ; it can be assumed that the variable y is not free in t') $= u'[t'/x]$.
Q.E.D.

Now the proof of the Church-Rosser property for ρ can be completed. So we assume that $t_0\ \rho\ t_1$, $t_0\ \rho\ t_2$, and we look for a term t_3 such that $t_1\ \rho\ t_3$, $t_2\ \rho\ t_3$. The proof is by induction on the length of t_0.
1) If t_0 has length 1, then it is a variable ; hence, by lemma 11, $t_0 = t_1 = t_2$; take $t_3 = t_0$.
2) If $t_0 = \lambda x\ u_0$, then, since $t_0\ \rho\ t_1$, $t_0\ \rho\ t_2$, by lemma 11 : $t_1 \equiv \lambda x\ u_1$, $t_2 \equiv \lambda x\ u_2$, and $u_0\ \rho\ u_1$, $u_0\ \rho\ u_2$. By induction hypothesis, $u_1\ \rho\ u_3$ and $u_2\ \rho\ u_3$ hold for some term u_3. Hence it is sufficient to take $t_3 = \lambda x\ u_3$.
3) If $t_0 = (v_0)u_0$, then, since $t_0\ \rho\ t_1$, $t_0\ \rho\ t_2$, by lemma 11, the different possible cases are :

a) $t_1 \equiv (v_1)u_1$, $t_2 \equiv (v_2)u_2$ with $u_0\ \rho\ u_1$, $v_0\ \rho\ v_1$, $u_0\ \rho\ u_2$, $v_0\ \rho\ v_2$. By induction hypothesis, $u_1\ \rho\ u_3$, $u_2\ \rho\ u_3$, $v_1\ \rho\ v_3$, $v_2\ \rho\ v_3$ hold for some u_3 and v_3. Hence it is sufficient to take $t_3 = (v_3)u_3$.

b) $t_1 \equiv (v_1)u_1$, with $u_0 \,\rho\, u_1$, $v_0 \,\rho\, v_1$; $v_0 \equiv \lambda x\, w_0$, $t_2 \equiv w_2[u_2/x]$, with $u_0 \,\rho\, u_2$, $w_0 \,\rho\, w_2$. Since $v_0 \,\rho\, v_1$, by lemma 11, $v_1 \equiv \lambda x\, w_1$, for some w_1 such that $w_0 \,\rho\, w_1$. Thus $t_1 \equiv (\lambda x\, w_1)u_1$. Since $u_0 \,\rho\, u_1$, $u_0 \,\rho\, u_2$, and $w_0 \,\rho\, w_1$, $w_0 \,\rho\, w_2$, by induction hypothesis, $u_1 \,\rho\, u_3$, $u_2 \,\rho\, u_3$, and $w_1 \,\rho\, w_3$, $w_2 \,\rho\, w_3$ hold for some u_3 and w_3. Hence, by rule 4, $(\lambda x\, w_1)u_1 \,\rho\, w_3[u_3/x]$, that is $t_1 \,\rho\, w_3[u_3/x]$. Now by lemma 12, $w_2[u_2/x] \,\rho\, w_3[u_3/x]$. So we obtain the expected result by taking $t_3 = w_3[u_3/x]$.

c) $v_0 \equiv \lambda x\, w_0$, $t_1 \equiv w_1[u_1/x]$, $t_2 \equiv w_2[u_2/x]$, with $u_0 \,\rho\, u_1$, $u_0 \,\rho\, u_2$, $w_0 \,\rho\, w_1$, $w_0 \,\rho\, w_2$. By induction hypothesis, $u_1 \,\rho\, u_3$, $u_2 \,\rho\, u_3$, $w_1 \,\rho\, w_3$, $w_2 \,\rho\, w_3$ hold for some u_3 and w_3. Hence, by lemma 12, $w_1[u_1/x] \,\rho\, w_3[u_3/x]$, $w_2[u_2/x] \,\rho\, w_3[u_3/x]$, that is to say $t_1 \,\rho\, w_3[u_3/x]$, $t_2 \,\rho\, w_3[u_3/x]$. The result follows by taking $t_3 = w_3[u_3/x]$.
Q.E.D.

REMARK. The intuitive meaning of the relation ρ is the following : $t \,\rho\, t'$ holds if and only if t' is obtained from t by contracting several redexes *occurring in t*.
For example, $(\lambda x(x)x)\lambda x\, x \,\rho\, (\lambda x\, x)\lambda x\, x$; a new redex has been created, but it cannot be contracted ; $(\lambda x(x)x)\lambda x\, x \,\rho\, \lambda x\, x$ does not hold.
In other words, $t \,\rho\, t'$ means that t and t' are constructed simultaneously : for t the steps of the construction are those described in the definition of terms, while for t', the same rules are applied, except that the following alternative is allowed : whenever $t = (\lambda x\, v)u$, t' can be taken either as $(\lambda x\, v')u'$ or as $v'[u'/x]$. This is what lemma 11 expresses.

β-EQUIVALENCE

The *β-equivalence* (denoted by $\simeq_\beta$) is defined as the least equivalence relation which contains β_0 (or β, which comes to the same thing). In other words :
$t \simeq_\beta t' \Leftrightarrow$ there exists a sequence $(t_1 = t), t_2, \ldots, t_{n-1}, (t_n = t')$, such that $t_i \,\beta_0\, t_{i+1}$ or $t_{i+1} \,\beta_0\, t_i$ for $1 \leq i < n$.
$t \simeq_\beta t'$ should be read as : t is β-equivalent to t'.
It follows from Church-Rosser's theorem that :
$t \simeq_\beta t' \Leftrightarrow t \,\beta\, u$ and $t' \,\beta\, u$ for some term u.
The side $\Leftarrow$ is obvious. For the purpose of proving $\Rightarrow$, consider the relation $\simeq$ defined by :
$t \simeq t' \Leftrightarrow t \,\beta\, u$ and $t' \,\beta\, u$ for some term u.
This relation contains β, and is reflexive and symmetric. It is also transitive, for if $t \simeq t'$, $t' \simeq t''$, then $t \,\beta\, u$, $t' \,\beta\, u$, and $t' \,\beta\, v$, $t'' \,\beta\, v$ for suitable u and v. By Church-Rosser's theorem, $u \,\beta\, w$ and $v \,\beta\, w$ hold for some term w ; thus $t \,\beta\, w$, $t'' \,\beta\, w$.
Hence $\simeq$ is an equivalence relation which contains β, so it also contains $\simeq_\beta$.
Q.E.D.

Therefore, a non-normalizable term cannot be β-equivalent to a normal term.

4. Eta-conversion

Proposition 7. **Assume that x is not free in t, nor x' in t'. If $\lambda x(t)x \equiv \lambda x'(t')x'$, then $t \equiv t'$.**

By definition of $\equiv$, if $\lambda x(t)x \equiv \lambda x'(t')x'$, then $((t)x)<y/x> \equiv ((t')x')<y/x'>$ for all variables y except a finite number ; this can be written $(t)y \equiv (t')y$ since x (resp. x') is not free in t (resp. t'). By definition of $\equiv$ again, $(t)y \equiv (t')y \Rightarrow t \equiv t'$.
Q.E.D.

A term of the form $\lambda x(t)x$ (where x is not free in t) is called an *η-redex*, its *contractum* being t. The previous proposition shows that this notion is defined on Λ.

A term of either of the forms $(\lambda x\, u)t$, $\lambda y(v)y$ (where y is not free in v) will be called a *$\beta\eta$-redex*.

We now define a binary relation η_0 on Λ ; $t \,\eta_0\, t'$ should be read as "t' is obtained by contracting an η-redex (or by an *η-reduction*) in the term t ". The definition is given by induction on t, as for β_0 :

if t is a variable, then there is no t' such that $t \,\eta_0\, t'$;
if $t = \lambda x\, u$, then $t \,\eta_0\, t'$ if and only if :
  either $t' = \lambda x\, u'$, with $u \,\eta_0\, u'$;
  or $u = (t')x$, with x not free in t' ;
if $t = (v)u$, then $t \,\eta_0\, t'$ if and only if :
  either $t' = (v)u'$ with $u \,\eta_0\, u'$,
  or $t' = (v')u$ with $v \,\eta_0\, v'$.

The relation $t \,\beta\eta_0\, t'$ (which means : "t' is obtained from t by contracting a $\beta\eta$-redex") is defined as : $t \,\beta_0\, t'$ or $t \,\eta_0\, t'$.
The *η-conversion* (resp. the *$\beta\eta$-conversion*) is defined as the least binary relation η (resp. $\beta\eta$) on Λ which is reflexive, transitive, and contains η_0 (resp. $\beta\eta_0$).

Proposition 8. **The $\beta\eta$-conversion is the least transitive λ-compatible binary relation $\beta\eta$ such that $(\lambda x\, u)t \,\beta\eta\, u[t/x]$ and $\lambda y(v)y \,\beta\eta\, v$ whenever y is not free in v.**

The proof is similar to that of proposition 4 (which is the analogue for β).

It can be proved, as for β, that $\beta\eta$ is the transitive closure of the binary relation ρ defined on L by : $u \, \rho \, u' \Leftrightarrow$ there exist a term v, and redexes $t_1,\ldots,t_k$ with contractums $t'_1,\ldots,t'_k$ such that $u \equiv v<t_1/x_1,\ldots,t_k/x_k>$, $u' \equiv v<t'_1/x_1,\ldots,t'_k/x_k>$.

Similarly : if $t \, \beta\eta \, t'$, then every free variable in t' is also free in t.

Proposition 9. If $t \, \beta\eta \, t'$ and $u \, \beta\eta \, u'$, then $u[t/x] \, \beta\eta \, u'[t'/x]$.

If $u = u'$, then the result follows from lemma 4, since the relation $\beta\eta$ is λ-compatible. It is enough to show now that $t \, \beta\eta \, t'$ and $u \, \beta\eta_0 \, u' \Rightarrow u[t/x] \, \beta\eta \, u'[t'/x]$.
The proof is by induction on the length of u. According to the definition of $\beta\eta_0$, the different possibilities for u,u' are :

i) $u = \lambda y \, v$, $u' = \lambda y \, v'$, and $v \, \beta\eta_0 \, v'$.
ii) $u = (w)v$ and $u' = (w)v'$, with $v \, \beta\eta_0 \, v'$.
iii) $u = (w)v$ and $u' = (w')v$, with $w \, \beta\eta_0 \, w'$.
iv) $u = (\lambda y \, w)v$ and $u' = w[v/y]$.
v) $u = \lambda y(u')y$, with y not free in u'.

Cases i) to iv) are settled exactly as in proposition 6. In case v), a representative of u is chosen of which no bound variable occurs in x,t,u'. So y is not a free variable in t and $y \neq x$. Thus $u[t/x] = \lambda y(u'[t/x])y$ and y does not occur in $u'[t/x]$. Hence $u[t/x] \, \beta\eta_0 \, u'[t/x]$. Moreover, $u'[t/x] \, \beta\eta \, u'[t'/x]$ (lemma 4), since $t \, \beta\eta \, t'$ and $\beta\eta$ is λ-compatible. Therefore $u[t/x] \, \beta\eta \, u'[t'/x]$.
Q.E.D.
This proposition is also an immediate consequence of lemma 14.

A term t is said to be *$\beta\eta$-normal* if it contains no $\beta\eta$-redex.
So the $\beta\eta$-normal terms are those obtained by applying, a finite number of times, the following rules :

any variable x is a $\beta\eta$-normal term ;
whenever t is $\beta\eta$-normal, then so is $\lambda x \, t$, except if $t = (u)x$, with x not free in u ;
whenever t,u are $\beta\eta$-normal, then so is (u)t, except if the first symbol in u is λ.

Theorem. The $\beta\eta$-conversion satisfies the Church-Rosser property.

The proof is on the same lines as for the β-conversion. Here ρ is defined as the least λ-compatible binary relation on Λ such that :
$t \, \rho \, t'$, $u \, \rho \, u' \Rightarrow (\lambda x \, u)t \, \rho \, u'[t'/x]$;
$t \, \rho \, t' \Rightarrow \lambda x(t)x \, \rho \, t'$ whenever x is not free in t.
The first thing to be proved is : $\beta\eta \supset \rho$.

For that purpose, note that $t \, \beta\eta \, t'$, $u \, \beta\eta \, u' \Rightarrow (\lambda x \, u)t \, \beta\eta \, u'[t'/x]$; indeed, since $\beta\eta$ is λ-compatible, we have $(\lambda x \, u)t \, \beta\eta \, (\lambda x \, u')t'$ and, on the other hand, $(\lambda x \, u')t' \, \beta\eta \, u'[t'/x]$; the result then follows, by transitivity.
Now $t \, \beta\eta \, t' \Rightarrow \lambda x(t)x \, \beta\eta \, t'$; this is immediate, by transitivity, since $\lambda x(t)x \, \beta\eta \, t$.
Therefore $\beta\eta$ is the transitive closure of ρ. It thus remains to prove that ρ satisfies the Church-Rosser property.
By definition, ρ is the set of all pairs of (equivalence classes of) terms obtained by applying, a finite number of times, the following rules :

1) $t \, \rho \, t$;
2) $t \, \rho \, t' \Rightarrow \lambda x \, t \, \rho \, \lambda x \, t'$;
3) $t \, \rho \, t'$ and $u \, \rho \, u' \Rightarrow (u)t \, \rho \, (u')t'$;
4) $t \, \rho \, t'$, $u \, \rho \, u' \Rightarrow (\lambda x \, u)t \, \rho \, u'[t'/x]$.
5) $t \, \rho \, t' \Rightarrow \lambda x(t)x \, \rho \, t'$ whenever x is not free in t.

The following lemmas are the analogues of lemmas 11 and 12.

Lemma 13. i) If $x \, \rho \, t'$, where x is a variable, then $t' = x$.
ii) If $\lambda x \, u \, \rho \, t'$, then either $t' \equiv \lambda x \, u'$ and $u \, \rho \, u'$, or $u \equiv (t)x$ and $t \, \rho \, t'$, with x not free in t.
iii) If $(v)u \, \rho \, t'$, then either $t' \equiv (v')u'$ with $u \, \rho \, u'$ and $v \, \rho \, v'$, or $v \equiv \lambda x \, w$ and $t' \equiv w'[u'/x]$ with $u \, \rho \, u'$ and $w \, \rho \, w'$.

Same proof as for lemma 11.

Lemma 14. Whenever $t \, \rho \, t'$ and $u \, \rho \, u'$, then $u[t/x] \, \rho \, u'[t'/x]$.

The proof proceeds by induction on the length of the derivation of $u \, \rho \, u'$ by means of rules 1 through 5 ; consider the last rule used :

if it is one of rules 1,2,3,4, then the proof is the same as in lemma 12 ;
if it is rule 5, then $u \equiv \lambda y(v)y$ and $v \, \rho \, u'$, with y not free in v. With a suitable choice of the representative of u, it can be assumed that y is not free in t. By induction hypothesis, $v[t/x] \, \rho \, u'[t'/x]$, then, by applying rule 5, we obtain $\lambda y(v[t/x])y \, \rho \, u'[t'/x]$ (since y is not free in $v[t/x]$), that is $u[t/x] \, \rho \, u'[t'/x]$.

Q.E.D.

Now the proof of the Church-Rosser property for ρ can be completed. So we assume that $t_0 \, \rho \, t_1$, $t_0 \, \rho \, t_2$, and we look for a term t_3 such that $t_1 \, \rho \, t_3$, $t_2 \, \rho \, t_3$. The proof is by induction on the length of t_0.
1) If t_0 has length 1, then it is a variable ; hence, by lemma 13, $t_0 = t_1 = t_2$; take $t_3 = t_0$.

2) If $t_0 = \lambda x\, u_0$, then, since $t_0\, \rho\, t_1$, $t_0\, \rho\, t_2$, by lemma 13, the different possible cases are :

a) $t_1 \equiv \lambda x\, u_1$, $t_2 \equiv \lambda x\, u_2$, and $u_0\, \rho\, u_1$, $u_0\, \rho\, u_2$. By induction hypothesis, $u_1\, \rho\, u_3$ and $u_2\, \rho\, u_3$ hold for some term u_3. Then it is sufficient to take $t_3 = \lambda x\, u_3$.

b) $t_1 \equiv \lambda x\, u_1$, and $u_0\, \rho\, u_1$; $u_0 \equiv (t'_0)x$, with x not free in t'_0, and $t'_0\, \rho\, t_2$.
According to lemma 13, since $u_0\, \rho\, u_1$ and $u_0 \equiv (t'_0)x$, there are two possibilities for u_1 :

i) $u_1 \equiv (t'_1)x$, with $t'_0\, \rho\, t'_1$. Now $t'_0\, \rho\, t_2$, thus, by induction hypothesis, $t'_1\, \rho\, t_3$ and $t_2\, \rho\, t_3$ hold for some term t_3. Note that, since $t'_0\, \rho\, t'_1$, all free variables in t'_1 are also free in t'_0, so x is not free in t'_1. Hence, by rule 5, $\lambda x(t'_1)x\, \rho\, t_3$, that is $t_1\, \rho\, t_3$.

ii) $t'_0 \equiv \lambda y\, u'_0$, $u_1 \equiv u'_1[x/y]$ and $u'_0\, \rho\, u'_1$. Since ρ is λ-compatible, $\lambda y\, u'_0\, \rho\, \lambda y\, u'_1$, that is $t'_0\, \rho\, \lambda y\, u'_1$. Now x is not free in t'_0 , hence it is neither free in u'_0 (provided that a representative of t'_0 has been chosen in which x is not bound). Since $u'_0\, \rho\, u'_1$, x is not free in u'_1, hence (lemma 5) $\lambda y\, u'_1 \equiv \lambda x\, u'_1[x/y]$, in other words $\lambda y\, u'_1 \equiv \lambda x\, u_1 \equiv t_1$. Now $t'_0\, \rho\, t_1$ because $t'_0\, \rho\, \lambda y\, u'_1$; and, since $t'_0\, \rho\, t_2$, there exists, by induction hypothesis, a term t_3 such that $t_1\, \rho\, t_3$, $t_2\, \rho\, t_3$.

c) $u_0 \equiv (t'_0)x$, with x not free in t'_0, and $t'_0\, \rho\, t_1$, $t'_0\, \rho\, t_2$. The conclusion follows immediately from the induction hypothesis, since t'_0 is shorter than t_0.

3) If $t_0 = (v_0)u_0$, then, since $t_0\, \rho\, t_1$, $t_0\, \rho\, t_2$, by lemma 13, the different possible cases are :

a) $t_1 \equiv (v_1)u_1$, $t_2 \equiv (v_2)u_2$ with $u_0\, \rho\, u_1$, $v_0\, \rho\, v_1$, $u_0\, \rho\, u_2$, $v_0\, \rho\, v_2$. By induction hypothesis, $u_1\, \rho\, u_3$, $u_2\, \rho\, u_3$, $v_1\, \rho\, v_3$, $v_2\, \rho\, v_3$ hold for some u_3 and v_3. Then it is sufficient to take $t_3 = (v_3)u_3$.

b) $t_1 \equiv (v_1)u_1$, with $u_0\, \rho\, u_1$, $v_0\, \rho\, v_1$; $v_0 \equiv \lambda x\, w_0$, $t_2 \equiv w_2[u_2/x]$, with $u_0\, \rho\, u_2$, $w_0\, \rho\, w_2$.
Since $v_0\, \rho\, v_1$, and $v_0 \equiv \lambda x\, w_0$, by lemma 13, the different possible cases are :

i) $v_1 \equiv \lambda x\, w_1$, with $w_0\, \rho\, w_1$. Then $t_1 \equiv (\lambda x\, w_1)u_1$.
Since $u_0\, \rho\, u_1$, $u_0\, \rho\, u_2$, and $w_0\, \rho\, w_1$, $w_0\, \rho\, w_2$, by induction hypothesis, $u_1\, \rho\, u_3$, $u_2\, \rho\, u_3$, and $w_1\, \rho\, w_3$, $w_2\, \rho\, w_3$ hold for some u_3, w_3. Thus, by rule 4, $(\lambda x\, w_1)u_1\, \rho\, w_3[u_3/x]$, that is $t_1\, \rho\, w_3[u_3/x]$. Hence, by lemma 14, $w_2[u_2/x]\, \rho\, w_3[u_3/x]$. The expected result is then obtained by taking $t_3 = w_3[u_3/x]$.

ii) $w_0 = (v'_0)x$, with x not free in v'_0, and $v'_0\, \rho\, v_1$. Then $(v'_0)x\, \rho\, w_2$; since $u_0\, \rho\, u_2$, it follows from lemma 14 that $((v'_0)x)[u_0/x]\, \rho\, w_2[u_2/x]$. But x is not free in v'_0, so this is equivalent to $(v'_0)u_0\, \rho\, t_2$.
Now $v'_0\, \rho\, v_1$ and $u_0\, \rho\, u_1$. Thus $(v'_0)u_0\, \rho\, (v_1)u_1$, in other words : $(v'_0)u_0\, \rho\, t_1$. Since $(v'_0)u_0$ is shorter than t_0 (because $v_0 \equiv \lambda x(v'_0)x$), there exists, by induction hypothesis, a term t_3 such that $t_1\, \rho\, t_3$, $t_2\, \rho\, t_3$.

c) $v_0 \equiv \lambda x\, w_0$, $t_1 \equiv w_1[u_1/x]$, $t_2 \equiv w_2[u_2/x]$, with $u_0\, \rho\, u_1$, $u_0\, \rho\, u_2$, $w_0\, \rho\, w_1$, $w_0\, \rho\, w_2$.
By induction hypothesis, $u_1\, \rho\, u_3$, $u_2\, \rho\, u_3$, $w_1\, \rho\, w_3$, $w_2\, \rho\, w_3$ hold for some u_3 and w_3.
Thus, by lemma 14, $w_1[u_1/x]\, \rho\, w_3[u_3/x]$, $w_2[u_2/x]\, \rho\, w_3[u_3/x]$, that is to say $t_1\, \rho\, w_3[u_3/x]$, $t_2\, \rho\, w_3[u_3/x]$. The result follows by taking $t_3 = w_3[u_3/x]$.
Q.E.D.

The ***βη-equivalence*** (denoted by $\simeq_{\beta\eta}$) is defined as the least equivalence relation which contains $\beta\eta$. In other words :

$t \simeq_{\beta\eta} t' \Leftrightarrow$ there exists a sequence $t = t_1, t_2, \ldots, t_{n-1}, t_n = t'$, such that either $t_i\, \beta\eta\, t_{i+1}$ or $t_{i+1}\, \beta\eta\, t_i$, for $1 \leq i < n$.

As for the β-equivalence, it follows from Church-Rosser's theorem that :

$t \simeq_{\beta\eta} t' \Leftrightarrow t\, \beta\eta\, u$ and $t'\, \beta\eta\, u$ for some term u.

The relation $\simeq_{\beta\eta}$ satisfies the "extensionality axiom", that is to say :

if $(t)u \simeq_{\beta\eta} (t')u$ holds for all u, then $t \simeq_{\beta\eta} t'$.

Indeed, it is enough to take u as a variable x which does not occur in t,t'. Since $\simeq_{\beta\eta}$ is λ-compatible, we have $\lambda x(t)x \simeq_{\beta\eta} \lambda x(t')x$; therefore, by η-reduction, $t \simeq_{\beta\eta} t'$.

REFERENCES FOR CHAPTER I.

[Bar84], [Chu41], [Hin86].
(The references are in the bibliography at the end of the book).

Chapter II

REPRESENTATION OF RECURSIVE FUNCTIONS

1. Head normal forms

In every λ-term, each subsequence of the form "$(\lambda$" corresponds to a unique redex (this is obvious since redexes are terms of the form $(\lambda x\ u)t$). This allows us to define, in any non-normal term t, the "leftmost redex in t". Let t' be the term obtained from t by contracting that leftmost redex : we say that t' is obtained from t by a ***leftmost β-reduction.***
Let t be an arbitrary λ-term. With t we associate a (finite or infinite) sequence of terms $t_0,t_1,...,t_n,...$ such that $t_0 = t$, and t_{n+1} is obtained from t_n by a leftmost β-reduction (if t_n is normal, then the sequence ends with t_n). We call it "the sequence obtained from t by leftmost β-reduction" (it is uniquely determined by t).
The following theorem will be proved in Chapter IV :

Theorem 1. If t is a normalizable term, then the sequence obtained from t by leftmost β-reduction ends with the normal form of t.

We see that this theorem provides a "normalizing strategy", which can be used for any normalizable term.

NOTATION. We will write $t \succ\succ u$ whenever u is obtained from t by a sequence of leftmost β-reductions.

The next proposition is simply a remark about the form of the λ-terms :

Proposition 1. Every term of the λ-calculus can be written, in a unique way, in the form $\lambda x_1...\lambda x_m(v)t_1...t_n$, where $x_1,...,x_m$ are variables, $t_1,...,t_n$ are terms $(m,n \geq 0)$, and v is either i) a variable or ii) a redex ($v = (\lambda x\ u)t$).

Recall that $(v)t_1...t_n$ denotes the term $(...((v)t_1)...)t_n$.

We prove the proposition by induction on the length of the considered term τ : the result is clear if τ is a variable.
If $\tau = \lambda x\ \tau'$, then τ' is determined by τ, and can be written in a unique way in the indicated form, by induction hypothesis ; thus the same holds for τ.
If $\tau = (w)v$, then v and w are determined by τ. If w starts with λ, then τ is a redex, so it is of the second form, and not of the first one. If w does not start with λ, then, by induction hypothesis, $w = (w')t_1...t_n$, where w' is a variable or a redex ; thus $\tau = (w')t_1...t_nv$, which is in one and only one of the indicated forms.
Q.E.D.

DEFINITIONS. A term τ is ***a head normal form*** (or ***in head normal form***) if it is of the first form indicated in proposition 1, namely if $\tau = \lambda x_1...\lambda x_m(x)t_1...t_n$, where x is a variable.
In the second case, if $\tau = \lambda x_1...\lambda x_m(\lambda x\ u)tt_1...t_n$, then the redex $(\lambda x\ u)t$ is called the ***head redex*** of τ.
The head redex of a term τ, when it exists (namely when τ is not a head normal form), is clearly the leftmost redex in τ.

It follows from proposition 1 that a term t is normal if and only if it is a head normal form : $\tau = \lambda x_1...\lambda x_m(x)t_1...t_n$, where $t_1,...,t_n$ are normal terms. In other words, a term is normal if and only if it is "hereditarily in head normal form".

The ***head reduction*** of a term τ is the (finite or infinite) sequence of terms $\tau_0,\tau_1,...,\tau_n,...$ such that $\tau_0 = \tau$, and τ_{n+1} is obtained from τ_n by a β-reduction of the head redex of τ_n if such a redex exists ; if not, τ_n is in head normal form, and the sequence ends with τ_n.

NOTATION. We will write $t \succ u$ whenever u is obtained from t by a sequence of head β-reductions.

A λ-term t is said to be ***solvable*** if, for any term u, there exist variables $x_1,...,x_k$ and terms $u_1,...,u_k,v_1,...,v_l$ $(k,l \geq 0)$ such that :
(i) $(t[u_1/x_1,...,u_k/x_k])v_1...v_l \simeq_\beta u$.

We have the following equivalent definitions :
(ii) t is solvable if and only if there exist variables $x_1,...,x_k$ and terms $u_1,...,u_k$, $v_1,...,v_l$ such that $(t[u_1/x_1,...,u_k/x_k])v_1...v_l \simeq_\beta I$ (I is the term $\lambda x\ x$).
(iii) t is solvable if and only if, given any variable x which does not occur in t, there exist terms $u_1,...,u_k$, $v_1,...,v_l$ such that $(t[u_1/x_1,...,u_k/x_k])v_1...v_l \simeq_\beta x$.

Obviously, (i) ⇒ (ii) ⇒ (iii). Now if $(t[u_1/x_1,\dots,u_k/x_k])v_1\dots v_l \simeq_\beta x$, then : $(t[u_1/x_1,\dots,u_k/x_k][u/x])v'_1\dots v'_l \simeq_\beta u$, and therefore $(t[u'_1/x_1,\dots,u'_k/x_k])v'_1\dots v'_l \simeq_\beta u$, where $u'_i = u_i[u/x]$, $v'_j = v_j[u/x]$; so we also have (iii) ⇒ (i).

REMARKS.

The following properties are immediate :

1) Let t be a closed term. Then t is solvable if and only if there exist terms $v_1,\dots,v_l$ such that $(t)v_1\dots v_l \simeq_\beta I$.

2) A term t is solvable if and only if its closure $\bar{t}$ is solvable (the closure of t is, by definition, the term $\bar{t} = \lambda x_1\dots\lambda x_n t$, where $x_1,\dots,x_n$ are the free variables occurring in t).

3) If (t)v is a solvable term, then t is solvable.

4) Of course, the head normal form of a term need not be unique. Nevertheless :
If a term t has a head normal form $t_0 = \lambda x_1\dots\lambda x_k(x)u_1\dots u_n$, then any head normal form of t can be written $\lambda x_1\dots\lambda x_k(x)u'_1\dots u'_n$, with $u_i \simeq_\beta u'_i$.

Indeed, let $t_1 = \lambda y_1\dots\lambda y_l(y)v_1\dots v_p$ be another head normal form of t. By the Church-Rosser theorem, there exists a term t_2 which can be obtained by β-reduction from t_0 as well as from t_1. Now, in t_0 (resp. t_1) all possible β-reductions have to be made in $u_1,\dots,u_n$ (resp. $v_1,\dots,v_p$). Hence $t_2 \equiv \lambda x_1\dots\lambda x_k(x)u'_1\dots u'_n \equiv \lambda y_1\dots\lambda y_l(y)v'_1\dots v'_p$, with $u_i\ \beta\ u'_i$, $v_j\ \beta\ v'_j$. This yields the expected result.

The following theorem will be proved in Chapter IV (th. 3, p. 53) :

Theorem 2. For every λ-term t, the following conditions are equivalent :
i) t is solvable ;
ii) t is β-equivalent to a head normal form ;
iii) the head reduction of t ends (with a head normal form).

2. Representable Functions

We define the *Booleans* $0 = \lambda x\lambda y\, y$ and $1 = \lambda x\lambda y\, x$. Then, for all terms t,u, ((0)t)u can be reduced (leftwards) to u , while ((1)t)u can be reduced to t.

Given two terms t,u and an integer k, let $(t)^k u$ denote the term (t)...(t)u (with k occurrences of t) ; in particular, $(t)^0 u = u$.

We define the term $\underline{k} = \lambda f\lambda x(f)^k x$; $\underline{k}$ is called "the numeral (or integer) k of the λ-calculus" (also known as ***Church numeral*** k, or ***Church integer*** k).
Notice that the Boolean 0 is the same term as the numeral $\underline{0}$, while the Boolean 1 is different from the numeral $\underline{1}$.

Let φ be a partial function defined on $\mathbb{N}^n$, with values either in ℕ or in {0,1}. Given a λ-term Φ, we say that Φ ***represents*** (resp. ***strongly represents***) the function φ if, for all $k_1,\dots,k_n \in \mathbb{N}$:

if $\varphi(k_1,\dots,k_n)$ is undefined, then $(\Phi)\underline{k_1}\dots\underline{k_n}$ is not normalizable (resp. not solvable) ;
if $\varphi(k_1,\dots,k_n) = k$, then $(\Phi)\underline{k_1}\dots\underline{k_n}$ is β-equivalent to $\underline{k}$ (or to k, in case the range of φ is {0,1}).

Clearly, for total functions, these two notions of representation are equivalent.

Theorem. Every partial recursive function from $\mathbb{N}^k$ to ℕ is (strongly) representable by a term of the λ-calculus.

Recall the definition of the class of partial recursive functions.
Given $f_1,\dots,f_k$, partial functions from $\mathbb{N}^n$ to ℕ, and g, partial function from $\mathbb{N}^k$ to ℕ, the partial function h, from $\mathbb{N}^n$ to ℕ, obtained by composition, is defined as follows :
$h(p_1,\dots,p_n) = g(f_1(p_1,\dots,p_n),\dots,f_k(p_1,\dots,p_n))$ if $f_1(p_1,\dots,p_n),\dots,f_k(p_1,\dots,p_n)$ are all defined, and $h(p_1,\dots,p_n)$ is undefined otherwise.
Let h be a partial function from ℕ to ℕ. If there exists an integer p such that $h(p) = 0$ and $h(q)$ is defined and different from 0 for all $q < p$, then we denote that integer p by $\mu n\{h(n) = 0\}$; otherwise $\mu n\{h(n) = 0\}$ is undefined.
We call ***minimization*** the operation which associates, with each partial function f from $\mathbb{N}^{k+1}$ to ℕ, the partial function g, from $\mathbb{N}^k$ to ℕ, such that $g(n_1,\dots,n_k) = \mu n\{f(n_1,\dots,n_k,n) = 0\}$.

The class of partial recursive functions is the least class of partial functions, with arguments and values in ℕ, closed under composition and minimization, and containing : the one argument constant function 0 and successor function ; the two arguments addition, multiplication, and characteristic function of the binary relation $x \le y$; and the projections P^k_n , defined by $P^k_n(x_1,\dots,x_n) = x_k$.

So it is sufficient to prove that the class of partial functions which are strongly representable by a term of the λ-calculus satisfies these properties.

The constant function 0 is represented by the term $\lambda d\, \underline{0}$.
The successor function on ℕ is represented by the term $suc = \lambda n\lambda f\lambda x((n)f)(f)x$.

The addition and the multiplication (functions from $\mathbb{N}^2$ to $\mathbb{N}$) are respectively represented by the terms $\lambda m\lambda n\lambda f\lambda x((m)f)((n)f)x$ and $\lambda m\lambda n\lambda f(m)(n)f$.
The characteristic function of the binary relation $m \leq n$ on $\mathbb{N}$ is represented by the term $M = \lambda m\lambda n(((m)A)\lambda d\ 1)((n)A)\lambda d\ 0$, where $A = \lambda f\lambda g(g)f$.
The function P^k_n is represented by the term $\lambda x_1 \ldots \lambda x_n\ x_k$.

From now on, we denote the term $(suc)^n\underline{0}$ by $\hat{n}$; so we have $\hat{n} \simeq_\beta \underline{n}$, and $(suc)\hat{n} = (n+1)\hat{}$.

REPRESENTATION OF COMPOSITE FUNCTIONS

Given any two λ-terms t,u, and a variable x with no free occurrence in t,u, the term $\lambda x(t)(u)x$ is denoted by $t{\circ}u$.

Lemma. $(\lambda g\ g{\circ}s)^k h \succ \lambda x(h)(s)^k x$ **for all closed terms s,h and every integer $k \geq 1$.**

Recall that $t \succ u$ means that u is obtained from t by a sequence of head β-reductions.
We prove the lemma by induction on k. The case $k = 1$ is clear. Assume the result for k ; then
$(\lambda g\ g{\circ}s)^{k+1}h = (\lambda g\ g{\circ}s)^k(\lambda g\ g{\circ}s)h \succ \lambda x((\lambda g\ g{\circ}s)h)(s)^k x$
(by induction hypothesis, applied with $(\lambda g\ g{\circ}s)h$ instead of h)
$\succ \lambda x(h{\circ}s)(s)^k x \equiv \lambda x(\lambda y(h)(s)y)(s)^k x \succ \lambda x(h)(s)^{k+1}x$.
Q.E.D.

Lemma. Let Φ,ν be two terms. Define : $<\Phi,\nu> = (((\nu)\lambda g\ g{\circ}suc)\Phi)\underline{0}$. Then :
if ν is not solvable, then neither is $<\Phi,\nu>$;
if $\nu \simeq_\beta n$ (Church numeral), then $<\Phi,\nu> \simeq_\beta (\Phi)n$; and if Φ is not solvable, then neither is $<\Phi,\nu>$.

The first statement follows from remark 3, p. 23.
If $\nu \simeq_\beta \underline{n}$, then : $(\nu)\lambda g\ g{\circ}suc \simeq_\beta (\underline{n})\lambda g\ g{\circ}suc = (\lambda f\lambda h(f)^n h)\lambda g\ g{\circ}suc \simeq_\beta \lambda h(\lambda g\ g{\circ}suc)^n h$;
by the previous lemma, this term gives, by head reduction, $\lambda h\lambda x(h)(suc)^n x$.
Hence $<\Phi,\nu> \simeq_\beta (\Phi)(suc)^n\underline{0} \simeq_\beta (\Phi)\underline{n}$. Therefore, if Φ is not solvable, then neither is $<\Phi,\nu>$ (remark 3, p. 23).
Q.E.D.

The term $<\Phi,\nu_1,\ldots,\nu_k>$ is defined, for $k \geq 2$, by induction on k :
$<\Phi,\nu_1,\ldots,\nu_k> = <<\Phi,\nu_1,\ldots,\nu_{k-1}>,\nu_k>$.

Lemma. Let $\Phi,\nu_1,\ldots,\nu_k$ **be terms such that each** ν_i **is either β-equivalent to a Church numeral, or not solvable. Then :**
if one of the ν_i's is not solvable, then neither is $<\Phi,\nu_1,\ldots,\nu_k>$;
if $\nu_i \simeq_\beta \underline{n}_i$ $(1 \leq i \leq k)$, then $<\Phi,\nu_1,\ldots,\nu_k> \simeq_\beta (\Phi)\underline{n}_1\ldots\underline{n}_k$.

The proof is by induction on k : let $\Psi = <\Phi,\nu_1,\ldots,\nu_{k-1}>$; then $<\Phi,\nu_1,\ldots,\nu_k> = <\Psi,\nu_k>$. If ν_k is not solvable, then, by the previous lemma, neither is $<\Psi,\nu_k>$. If ν_k is solvable (and β-equivalent to a Church numeral), and if one of the ν_i's $(1 \leq i \leq k-1)$ is not solvable, then Ψ is not solvable (induction hypothesis), and hence neither is $<\Psi,\nu_k>$ (previous lemma). Finally, if $\nu_i \simeq_\beta \underline{n}_i$ $(1 \leq i \leq k)$, then, by induction hypothesis, $\Psi \simeq_\beta (\Phi)\underline{n}_1\ldots\underline{n}_{k-1}$; therefore, $<\Psi,\nu_k> \simeq_\beta (\Phi)\underline{n}_1\ldots\underline{n}_k$ (previous lemma).
Q.E.D.

Proposition. Let $f_1,\ldots,f_k$ be partial functions from $\mathbb{N}^n$ to $\mathbb{N}$, and g a partial function from $\mathbb{N}^k$ to $\mathbb{N}$. Assume that these functions are all strongly representable by λ-terms ; then so is the composite function $g(f_1,\ldots,f_k)$.

Choose terms $\Phi_1,\ldots,\Phi_k,\Psi$ which strongly represent the functions $f_1,\ldots,f_k,g$, respectively. Then the term $\chi = \lambda x_1\ldots\lambda x_n<\Psi, (\Phi_1)x_1\ldots x_n,\ldots, (\Phi_k)x_1\ldots x_n>$ strongly represents the composite function $g(f_1,\ldots,f_k)$. Indeed, if $\underline{p}_1,\ldots,\underline{p}_n$ are Church numerals, then $(\chi)\underline{p}_1\ldots\underline{p}_n \simeq_\beta <\Psi, (\Phi_1)\underline{p}_1\ldots\underline{p}_n,\ldots, (\Phi_k)\underline{p}_1\ldots\underline{p}_n>$. Now each of the terms $(\Phi_i)\underline{p}_1\ldots\underline{p}_n$ is, either unsolvable (and in that case $f_i(p_1,\ldots,p_n)$ is undefined), or β-equivalent to a Church numeral $\underline{q}_i$ (then $f_i(p_1,\ldots,p_n) = q_i$). If one of the terms $(\Phi_i)\underline{p}_1\ldots\underline{p}_n$ is not solvable, then, by the previous lemma, neither is $(\chi)\underline{p}_1\ldots\underline{p}_n$. If $(\Phi_i)\underline{p}_1\ldots\underline{p}_n \simeq_\beta \underline{q}_i$ (where $1 \leq i \leq k$ and $\underline{q}_i$ is a Church numeral), then, by the previous lemma, we have $(\chi)\underline{p}_1\ldots\underline{p}_n \simeq_\beta (\Psi)\underline{q}_1\ldots\underline{q}_k$.
Q.E.D.

3. Fixed point combinators

A *fixed point combinator* is a closed term M such that $(M)F \simeq_\beta (F)(M)F$ for every term F. The main point is the existence of such terms. Here are two examples :

Let Y be the term $\lambda f\ (\lambda x(f)(x)x)\ \lambda x(f)(x)x$; then, for any term F, $(Y)F \simeq_\beta (F)(Y)F$.

Indeed, $(Y)F \succ (G)G$, where $G = \lambda x(F)(x)x$; therefore :
$(Y)F \succ (\lambda x(F)(x)x)G \succ (F)(G)G \simeq_\beta (F)(Y)F$.

Y is known as Curry's fixed point combinator. Note that we have neither $(Y)F \succ (F)(Y)F$, nor $(Y)F \; \beta \; (F)(Y)F$.

Now let Θ be the term $(A)A$, where $A \equiv \lambda a\lambda f(f)(a)af$. Then, for any term F, we have $(\Theta)F \succ (F)(\Theta)F$.

Indeed, $(\Theta)F \equiv (A)AF \succ (F)(A)AF \equiv (F)(\Theta)F$.

Θ is called Turing's fixed point combinator.

REPRESENTATION OF FUNCTIONS DEFINED BY MINIMIZATION

Lemma. **There exists a closed term Δ such that, for all terms Φ,n :**
$(\Delta)\Phi n \succ (\; ((\Phi)n)((\Delta)\Phi)(suc)n\;)n$.

Let $T = \lambda\delta\lambda\varphi\lambda\nu(\; ((\varphi)\nu)((\delta)\varphi)(suc)\nu\;)\nu$. Then Δ is defined as a fixed point of T, by means, for example, of Curry's fixed point combinator : we take $\Delta = (D)D$, where $D = \lambda x(T)(x)x$.
Then
$(\Delta)\Phi n = (D)D\Phi n \succ ((T)(D)D)\Phi n = (T)\Delta\Phi n \succ (\; ((\Phi)n)((\Delta)\Phi)(suc)n\;)n$.
Q.E.D.

Lemma. **Let b,t_0,t_1 be terms, and suppose $b \simeq_\beta 1$ (Boolean). Then $(b)t_0t_1 \succ t_0$.**

We first prove that $b \succ 1$; indeed, by theorem 2, we have $b \succ c$, where c is a head normal form, and $c \simeq_\beta 1$. Therefore, both 1 (which is $\lambda x\lambda y\; x$) and c are head normal forms for b ; hence, by remark 4, p. 23, we have $c = \lambda x\lambda y\; x$.
Now the proof of the lemma proceeds by induction on the length of the head reduction of b.
If this length is 0, then $b = 1$ and the result is immediate.
If b does not start with λ, and if b' is the term obtained from b by a head reduction, then $(b)t_0t_1 \succ (b')t_0t_1 \succ t_0$, by induction hypothesis.
If b starts with λ, then either $b = \lambda x(\lambda z\; v)uu_1...u_n$, or $b = \lambda x\lambda y(\lambda z\; v)uu_1...u_n$ (note that there are at most two occurrences of λ in a head position in b, since b reduces to $\lambda x\lambda y\; x$ by head reduction). Consider for example the second case. By a head reduction in b, we obtain $b' = \lambda x\lambda y(v[u/z])u_1...u_n$. It is then obvious that :
$(b')t_0t_1 \succ (v[u/z][t_0/x,t_1/y])u'_1...u'_n$, where $u'_i = u_i[t_0/x,t_1/y]$ (note that x,y do not occur in t_0,t_1).
Now, by induction hypothesis, $(b')t_0t_1 \succ t_0$. Therefore :
(*) $\quad (v[u/z][t_0/x,t_1/y])u'_1...u'_n \succ t_0$.

On the other hand, we have $(b)t_0t_1 \succ (\lambda z\; v[t_0/x,t_1/y])u'u'_1...u'_n$, where $u' = u[t_0/x,t_1/y]$. By another head reduction, we obtain :
$(b)t_0t_1 \succ (v[t_0/x,t_1/y][u'/z])u'_1...u'_n$.
But, according to lemma I.10, we have $v[t_0/x,t_1/y][u'/z] \equiv v[u/z][t_0/x,t_1/y]$. It follows then from (*) that $(b)t_0t_1 \succ t_0$.
Q.E.D.

A shorter proof of this lemma can be given, which uses results from Chapters III and IV :
We have $\vdash_{\mathcal{D}\Omega} 1 : X,Y \to X$, where X and Y are two type variables ; since $b \simeq_\beta 1$, theorem IV.1 shows that $\vdash_{\mathcal{D}\Omega} b : X,Y \to X$.
Let $\mathcal{I}$ be an interpretation, in the sense of Chapter III, such that :
$|X|_{\mathcal{I}} = \{\tau;\; \tau \succ t_0\}$ and $|Y|_{\mathcal{I}} = \{\tau;\; \tau \succ t_1\}$.
Since $\vdash_{\mathcal{D}\Omega} b : X,Y \to X$, it follows from the adequacy lemma (Ch. III) that $b \in |X,Y \to X|_{\mathcal{I}}$.
Now we have clearly $t_0 \in |X|_{\mathcal{I}}$ and $t_1 \in |Y|_{\mathcal{I}}$. Thus $(b)t_0t_1 \in |X|_{\mathcal{I}}$, which means that $(b)t_0t_1$ reduces to t_0 by head reduction.
Q.E.D.

Lemma. **Let Φ be a λ-term and $n \in \mathbb{N}$.**
If $(\Phi)\underline{n}$ is not solvable, then neither is $(\Delta)\Phi\underline{n}$.
If $(\Phi)\underline{n} \simeq_\beta 0$ (Boolean), then $(\Delta)\Phi\underline{n} \simeq_\beta \underline{n}$.
If $(\Phi)\underline{n} \simeq_\beta 1$ (Boolean), then $((\Delta)\Phi)\hat{n} \succ ((\Delta)\Phi)(suc)\hat{n}$.
(Recall that $\hat{n} = (suc)^n\underline{0}$).

Indeed, it follows from the lemma where Δ was defined that :
$(\Delta)\Phi\underline{n} \succ (\; ((\Phi)\underline{n})((\Delta)\Phi)(suc)\underline{n}\;)\underline{n}$. Hence, if $(\Phi)\underline{n}$ is not solvable, then neither is $(\Delta)\Phi\underline{n}$ (remark 3, p. 23). Obviously, if $(\Phi)\underline{n} \simeq_\beta 0$ (Boolean), then $(\Delta)\Phi\underline{n} \simeq_\beta \underline{n}$.
On the other hand, according to the same lemma, we also have :
$(\Delta)\Phi\hat{n} \succ (\; ((\Phi)\hat{n})((\Delta)\Phi)(suc)\hat{n}\;)\hat{n}$; by the previous lemma, if $(\Phi)\hat{n} \simeq_\beta 1$ (Boolean), then $(\; ((\Phi)\hat{n})((\Delta)\Phi)(suc)\hat{n}\;)\hat{n} \succ ((\Delta)\Phi)(suc)\hat{n}$.
Therefore $(\Delta)\Phi\hat{n} \succ ((\Delta)\Phi)(suc)\hat{n}$.
Q.E.D.

Proposition. **Let $f(n_1,...,n_k,n)$ be a partial function from $\mathbb{N}^{k+1}$ to $\mathbb{N}$, and suppose that it is strongly representable by a term of the λ-calculus. Then the partial function defined by $g(n_1,...,n_k) = \mu n\{f(n_1,...,n_k,n) = 0\}$ is also strongly representable.**

Let φ be the partial function from $\mathbb{N}^{k+1}$ to $\{0,1\}$, which has the same domain as f, and such that $\varphi(n_1,...,n_k,n) = 0 \Leftrightarrow f(n_1,...,n_k,n) = 0$. Then $g(n_1,...,n_k) = \mu n\{\varphi(n_1,...,n_k,n) = 0\}$.

Let F denote a λ-term which strongly represents f ; then the term :
$\Phi = \lambda n_1...\lambda n_k \lambda n(\Gamma)(F)n_1...n_k n$, where $\Gamma = \lambda n((n)\lambda d\, 1)0$, strongly represents φ (Γ represents the characteristic function of ℕ − {0}).
Now consider the term Δ constructed above. The term $G = \lambda n_1...\lambda n_k((\Delta)\Phi n_1...n_k)\underline{0}$ strongly represents the function g. Indeed :

if $g(n_1,...,n_k)$ is defined and equal to p, then $\varphi(n_1,...,n_k,n)$ is defined and equal to 1 for n < p and to 0 for n = p. Thus $(\Phi)\underline{n_1}...\underline{n_k}\underline{n} \simeq_\beta 1$ for n < p, and $(\Phi)\underline{n_1}...\underline{n_k}\underline{p} \simeq_\beta 0$. So we can successively deduce from the previous lemma (since $\underline{0} = \hat{0}$) :
$((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0} \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{1} \succ ... \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{p} \simeq_\beta \underline{p}$.

if $g(n_1,...,n_k)$ is undefined, there are two possibilities :
i) $\varphi(n_1,...,n_k,n)$ is defined and equal to 1 for n < p and is undefined for n = p. Then we can successively deduce from the previous lemma (since $\underline{0} = \hat{0}$) :
$((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0} \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{1} \succ ... \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{p}$; the last term obtained is not solvable, since neither is $\Phi\underline{n_1}...\underline{n_k}\underline{p}$ (previous lemma). Consequently, $((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0}$ is not solvable (theorem 2,iii) ;
ii) $\varphi(n_1,...,n_k,n)$ is defined and equal to 1 for all n. Then (again by the previous lemma) :
$((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0} \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{1} \succ ... \succ ((\Delta)\Phi\underline{n_1}...\underline{n_k})\hat{n} \succ ...$
So the head reduction of $((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0}$ does not end. Therefore, by theorem 2, $((\Delta)\Phi\underline{n_1}...\underline{n_k})\underline{0}$ is not solvable.
Q.E.D.

It is intuitively clear, according to Church's thesis, that any partial function from $ℕ^k$ to ℕ, which is representable by a λ-term, is partial recursive. We shall not give a formal proof of this fact. So we can state the

Church-Kleene theorem. The partial functions from $ℕ^k$ to ℕ which are representable (resp. strongly representable) by a term of the λ-calculus are the partial recursive functions.

4. The second fixed point theorem

Consider a recursive enumeration : $n \longmapsto t_n$ of the terms of the λ-calculus. The inverse function will be denoted by $t \longmapsto [\![t]\!]$: more precisely, if t is a λ-term, then $[\![t]\!]$ is the Church numeral $\underline{n}$ such that $t_n = t$, which will be called the numeral of t.
The function $n \longmapsto [\![(t_n)\underline{n}]\!]$ is thus recursive, from ℕ to ℕ. By the previous theorem, there exists a term δ such that $(\delta)\underline{n} \simeq_\beta [\![(t_n)\underline{n}]\!]$, for every integer n.

Now, given an arbitrary term F, let $B = \lambda n(F)(\delta)n$. Then, for any integer n, we have $(B)\underline{n} \simeq_\beta (F)[\![(t_n)\underline{n}]\!]$. Take $\underline{n} = [\![B]\!]$, that is to say $t_n = B$; then $(t_n)\underline{n} = (B)[\![B]\!]$. If we denote the term $(B)[\![B]\!]$ by A, we obtain $A \simeq_\beta (F)[\![A]\!]$. So we have proved the

Theorem. For every λ-term F, there exists a λ-term A such that $A \simeq_\beta (F)[\![A]\!]$.

Theorem. Let $\mathcal{X},\mathcal{Y}$ be two non-empty disjoint sets of terms, which are saturated under the equivalence relation $\simeq_\beta$. Then $\mathcal{X}$ and $\mathcal{Y}$ are recursively inseparable.

Suppose that $\mathcal{X}$ and $\mathcal{Y}$ are recursively separable. This means that there exists a recursive set $\mathcal{A} \subset \Lambda$ such that $\mathcal{X} \subset \mathcal{A}$ and $\mathcal{Y} \subset \mathcal{A}^c$ (the complement of $\mathcal{A}$). By assumption, there exist terms ξ and η such that $\xi \in \mathcal{X}$ and $\eta \in \mathcal{Y}$. Since the characteristic function of $\mathcal{A}$ is recursive, there is a term Θ such that, for every integer n : $(\Theta)\underline{n} \simeq_\beta 1 \Leftrightarrow t_n \in \mathcal{A}$ and $(\Theta)\underline{n} \simeq_\beta 0 \Leftrightarrow t_n \notin \mathcal{A}$.
Now let $F = \lambda n(\Theta)n\eta\xi$. According to the previous theorem, there exists a term A such that $(F)[\![A]\!] \simeq_\beta A$, which implies $(\Theta)[\![A]\!]\eta\xi \simeq_\beta A$.
If $A \in \mathcal{A}$, then, by the definition of Θ, $(\Theta)[\![A]\!] \simeq_\beta 1$, hence $(\Theta)[\![A]\!]\eta\xi \simeq_\beta \eta$. Therefore $A \simeq_\beta \eta$. Since $\eta \in \mathcal{Y} \subset \mathcal{A}^c$ and $\mathcal{Y}$ is saturated under the equivalence relation $\simeq_\beta$, we conclude that $A \in \mathcal{Y}$, thus $A \notin \mathcal{A}$, which is a contradiction.
Similarly, if $A \notin \mathcal{A}$, then $(\Theta)[\![A]\!] \simeq_\beta 0$, hence $(\Theta)[\![A]\!]\eta\xi \simeq_\beta \xi$, and $A \simeq_\beta \xi$. Since $\xi \in \mathcal{X} \subset \mathcal{A}$ and $\mathcal{X}$ is saturated under the equivalence relation $\simeq_\beta$, we conclude that $A \in \mathcal{X}$, thus $A \in \mathcal{A}$, which is again a contradiction.
Q.E.D.

Corollary. The set of normalizable (resp. solvable) terms of the λ-calculus is not recursive.

Apply the previous result : take $\mathcal{X}$ as the set of normalizable (resp. solvable) terms, and $\mathcal{Y} = \mathcal{X}^c$.

REFERENCES FOR CHAPTER II.

[Bar84], [Hin86].
(The references are in the bibliography at the end of the book).