

Contents

Abstract	iii
Acknowledgements	iv
Introduction	v
1 Lambda Calculus	1
1.1 λ -terms	1
1.2 Substitution	2
1.3 β -reduction	3
1.4 Head Normal Form	4
1.5 Reduction strategies	4
1.5.1 Normal Order Reduction	5
1.5.2 Call-by-Name Reduction	6
1.5.3 Call-by-Value Reduction	7
1.5.4 Head Reduction	8
1.5.5 Applicative Order Reduction	9
1.6 Comparisons between Different Reduction Strategies	10
1.7 λ -calculus with Explicit Substitution and Garbage Collection	10
1.7.1 The explicit substitution calculi	11
1.7.2 Reduction in λ_{xgc} -calculus	12
1.8 Normal β -reduction vs Explicit Substitution and garbage collection	12
1.8.1 $\xrightarrow{\beta}$ vs \xrightarrow{bx}	13
1.8.2 $\xrightarrow{\beta}$ vs \xrightarrow{bx} vs \xrightarrow{bxgc}	14
2 The Curry type assignment system	16
2.1 The System $\lambda \rightarrow$ -Curry	16
2.2 Properties of $\lambda \rightarrow$	18
2.3 The principal type property	18
3 A Purely-Functional Programming Language — Haskell	20
3.1 Functional Programming	20
3.2 Haskell	20
3.2.1 Pattern Matching	21
3.2.2 Guards	21
3.2.3 Local Definition	21
3.2.4 Self-defined Types	22
3.3 Model the λ -calculus and the Curry Type Assignment System in Haskell	22
3.3.1 Data Types	22
3.3.2 Parser	23
3.3.3 Main mechanism	24
3.4 Issues and Weakness of Implementation in Haskell	24

4	The λ-calculus and Curry Type Assignment System Implementation in Java	25
4.1	Different operational order between Haskell and Java	25
4.2	λ -calculus Representation	26
4.2.1	λ -terms	26
4.2.2	Interpreter	27
4.3	Curry Type Assignment System Representation	28
4.3.1	Types	28
4.3.2	Type substitution	29
4.3.3	Other data types	29
4.3.4	Core type assignment mechanism implementation	30
4.4	Explicit substitution and garbage collection	31
4.5	Syntax	31
5	GUI development	32
5.1	Layout	32
5.2	Functionalities	33
5.2.1	‘Normalize’ tracer	33
5.2.2	‘Trace’ tracer	34
5.2.3	‘Single-step’ tracer	34
5.2.4	Deal with non-terminating terms by trace	35
5.2.5	Explicit Substitution	36
6	Java vs Haskell, Pros and Cons	37
6.1	Pros	37
6.2	Cons	37
7	Conclusion	38

Abstract

Acknowledgements

Introduction

The lambda calculus can be called the smallest universal programming language of the world[10]. It is an example of formal system which consists a language of λ -terms and some auxiliary notions such as free variable and subterms, and the transformation theory. It was introduced by Alonzo Church in the 1930s as a way of formalizing the concepts of effective computability. Lambda calculus is a formalism that can be used to express and evaluate any executable function. It is thus equivalent to Turing machines. Most of the functional programming languages today, like Miranda, Haskell, are based on the lambda calculus. Imperative programming languages such as Fortran, Java, C are based on the way a Turing machine is instructed: by a sequence of statements.

There are typed and untyped variants of lambda calculus. In type lambda calculus, types are usually objects of a syntactic nature that are assigned to lambda term. A type system consists a set of rules that assign a property called a *type* to the various constructs — such as variable, expression, functions or modules. There are several reasons to add types to a program, one of the most important reason to add type properties to a program is to reduce bugs. It is possible to build an abstract interpretation of programs by treating terms as objects. The abstractions can be regarded as the interface of objects, by checking whether the interfaces have been connected properly we can decide whether the program is defined in a consistent way.

Because of the infinite explicit rewrite rules for β -reduction rule schema, the lambda calculus is not a comprehensive computational model in the sense that the number of β -reduction steps required to compute something does not correspond well to the number of computation steps needed by any reasonable ‘real’ computer[2]. A new lambda calculus with explicit substitution, λxgc , is introduced by Bloo and Rose[2] in 1995. It retains the use of ordinary variable names and includes reduction rules for explicit garbage collection. The lambda calculus with explicit substitution is a conservative extension which preserves strong normalisation of the untyped lambda calculus.

Implementing untyped lambda calculus in functional programming language is simple and straightforward, since those functional programming languages are built based on the lambda calculus. However, extend the untyped lambda calculus into typed is a matter of patient specification. This project firstly implements the untyped lambda calculus in Haskell, and extends it into typed variant by Curry type assignment system. Moreover, it also provides multiple contracting(or normalizing) paths of a λ -term: normal order, call-by-name, call-by-value, head reduction, and applicative order. Furthermore, the typed lambda calculus is then extended into typed lambda calculus with explicit substitution and garbage collection. However, since Haskell is ‘lazy’, the implementation is lack of usability as a tool. Therefore, we transfer the Haskell implementation into Java, with usable GUI.

Chapter 1

Lambda Calculus

The Lambda Calculus is an example of a formal system which consists a language of lambda terms and some auxiliary notions such as *free variables* and *subterms*, and the transformation theory. The core of the theory is the notion of substitution - the driving force behind function application.

1.1 λ -terms

The expressions in lambda calculus can be formalized into following:

Definition 1.1 (λ -TERMS) λ -terms can be defined by the following rules:

- a variable, x , itself is a lambda term
- if M is a lambda term, and x is a variable, then $(\lambda x.M)$ is a lambda term
- if M and N are both lambda terms, then (MN) is a lambda term

A lambda term is valid if and only if it can be obtained by repeated application of these rules. However, some parentheses can be omitted in certain forms. For example the leftmost, outermost brackets are usually omitted.

A lambda abstraction $\lambda x.M$ takes a single input x and returns a term M . Thus, it defines an **anonymous** function. For example $\lambda x.x^2$ is a lambda abstraction for the function $f(x) = x^2$ using the term x^2 for M , then we can say that $f = \lambda x.x^2$. The function is anonymous since we write $(x^2)2$ rather than $f3$.

An application (MN) applies the input term N to the function M . For example $(\lambda x.x^2)2$ is an application that applies 2 to the function $f(x) = x^2$, it return 2^2 which equals 4.

As mentioned above, leftmost, outermost brackets are usually omitted. Therefore, $M_1M_2(M_3M_4)$ stands for $((M_1 \cdot M_2)(M_3 \cdot M_4))$. Similarly, λ can be omitted in repeated abstractions, for example $\lambda x_1x_2.M$ stands for $\lambda x_1.(\lambda x_2.M)$.

Definition 1.2 (free and bound variables) The set of free and bound variables are defined inductively by the following function:

$$\begin{array}{llll} FV(x) & = & \{x\} & BV(x) & = & \emptyset \\ FV(\lambda x.M) & = & FV(M) \setminus \{x\} & BV(\lambda x.M) & = & BV(M) \cup \{x\} \\ FV(MN) & = & FV(M) \cup FV(N) & BV(MN) & = & BV(M) \cup BV(N) \end{array}$$

For example, the lambda term $\lambda x.x$ has no free variables but a bound variable x . The function $\lambda x.x + y$ only has a single free variable y and a bound variable x . Notice that, the sets of bound and free variables are not necessarily disjoint; x occurs both bound and free in:

$$x(\lambda xy.x)$$

1.2 Substitution

A cursory approach to define the substitution operation leads to the problem of *variable capture*. It occurs when we substitute a term containing a free variable into a scope where the variable becomes bound. For example:

$$(\lambda xy.xy)y \neq \lambda y.yy$$

The free occurrence of y in the left hand term becomes confused with the bound variable after substitution. To avoid *variable capture*, we define a capture-avoiding substitution. The notion $[x := N]$ indicates the replacement of a variable x by a term N :

- (1) $x[x := N] = N$
- (2) $y[x := N] = y,$ *if* $y \neq x$
- (3) $(\lambda y.M)[x := N] = \lambda y.M,$ *if* $y = x$
- (4) $(\lambda y.M)[x := N] = \lambda y.(M[x := N]),$ *if* $y \notin FV(N) \ \& \ y \neq x$
- (5) $(\lambda y.M)[x := N] = \lambda z.(M[y := z])[x := N],$ *if* $y \in FV(N) \ \& \ y \neq x, z \text{ new}$
- (6) $(M_1 M_2)[x := N] = (M_1[x := N])(M_2[x := N])$

Definition 1.3 (α -equivalent) M is α -equivalent to N , written $M \equiv_\alpha N$, if N results from M by a series of changes of bound variable.

The notion of alpha-equivalent (or congruence) is a basic form of equivalence defined in lambda calculus. It captures the property that the particular choice of a bound variable in a lambda abstraction does not matter. For instance, $\lambda x.x$ and $\lambda y.y$ are α -congruent lambda terms which represent the same function. Notice that, the term-variable x and y are not α -equivalent terms since they are not bound in a lambda abstraction.

With the property of alpha equivalence, we can define the α -conversion:

$$\lambda x.M =_\alpha \lambda z.(M[z := x])$$

In some sense, the α -conversion is defined by the spirit of the (5) substitution rule. If we have a function $\lambda x.M$, then we can apply a substitution to this function: $(\lambda x.M)[z := x]$. According to the fifth substitution rule, it is unfolded to $(\lambda z.M[z := x])$. Using the α -conversion, we can always rename the bound variables of a term.

Definition 1.4 (Variable Convention[7]) If M_1, \dots, M_n occur in a certain context then in these terms all bound variables are chosen to be different from free variables.

The alpha-conversion is built based on the variable convention. Alpha conversion is used to allow bound variable names to be changed. For example, alpha-conversion of $\lambda x.x$ might get $\lambda y.y$. Terms that differ only by alpha-conversion are called alpha-equivalence (or alpha-congruence) as mentioned in the definition 1.3. By the assumption that free and bound variables are always different (variable convention), and that alpha-conversion will take place whenever a variable is both free and bound. The definition of term substitution becomes:

$$\begin{aligned} x[x := N] &= N \\ y[x := N] &= y, & \text{if } y \neq x \\ (\lambda y.M)[x := N] &= \lambda y.(M[x := N]) \\ (M_1 M_2)[x := N] &= (M_1[x := N])(M_2[x := N]) \end{aligned}$$

By this definition of term substitution, the *variable capture* is avoided. Since that y is bound in $\lambda y.M$ of the context:

$$\lambda y.M[x := N]$$

y can only be bound in N according to variable convention rule, otherwise, if y is also free in N , it would be renamed by another variable.

Here is an example of its use:

$$\begin{aligned}
& (\lambda xyz.xzy)(\lambda xz.x) \\
= & \lambda yz.(\lambda xz.x)zy \\
= & \lambda yz.(\lambda xw.x)zy \text{ by the variable convention} \\
= & \lambda yz.(\lambda w.z)y \\
= & \lambda yz.z
\end{aligned}$$

1.3 β -reduction

There are various notions of reduction for λ -terms, but the principle one is β -reduction. β -reduction is the one-step reduction relation, written as $' \rightarrow'_\beta$.

Definition 1.5 (β -reduction) For λ -terms M and N , we say that M β -reduces in one step to N , written as $M \rightarrow_\beta N$:

$$(\lambda x.M)N \rightarrow_\beta M[x := N]$$

$$\begin{array}{c}
\frac{M \rightarrow_\beta N}{MZ \rightarrow_\beta NZ} \\
\frac{M \rightarrow_\beta N}{ZM \rightarrow_\beta ZM} \\
\frac{M \rightarrow_\beta N}{\lambda x.M \rightarrow_\beta \lambda x.N}
\end{array}$$

The reduction relation, written $' \rightarrow'_\beta$, is the reflexive, transitive closure of the one-step reduction relation. The one-step reduction relation allows a single step of reduction, while the reduction relation allows many steps. The reflexive transitive closure is defined as follows:

$$\begin{array}{c}
\frac{M \rightarrow_\beta N}{M \twoheadrightarrow_\beta N} \\
\frac{M \twoheadrightarrow_\beta N}{M \twoheadrightarrow_\beta M} \\
\frac{M \twoheadrightarrow_\beta N \quad N \twoheadrightarrow_\beta Z}{M \twoheadrightarrow_\beta Z}
\end{array}$$

For the notion $'M \rightarrow_\beta N'$, it is read as ' M β -reduces to N '. The first rule defines that the reduction relation is reflexive to one-step reduction relation, while the third rule indicates the transitivity of reduction relation.

Finally, the β -equality, written as $'=_\beta'$. For λ -terms A and B , we say that $A =_\beta B$ if either $A \equiv B$ or there exists a sequence of reduction starting with A , ending with B . It is the equivalence relation generated by \twoheadrightarrow_β :

$$\begin{array}{c}
\frac{M \twoheadrightarrow_\beta N}{M =_\beta N} \\
\frac{M =_\beta N}{N =_\beta M} \\
\frac{M =_\beta N \quad N =_\beta Z}{M =_\beta Z}
\end{array}$$

For the notion $'M =_\beta N'$, we say ' M β -equivalent to N '. Following is a simple example which applies β -reduction:

$$\begin{aligned}
& (\lambda x.x^2)3 \\
= & (x^2)[x := 3] \\
= & 3^2 \\
= & 9
\end{aligned}$$

Definition 1.6 (β -redex) A β -redex of a λ -term M is a subterm of M of the form $(\lambda x.P)Q$. A term M is called in β -normal form if it has no β -redex.

A β -redex is essentially a candidate for an application of β -reduction. A term M has a β -normal form if there exists a term N such that N is in β -normal form and $M \rightarrow_\beta N$.

1.4 Head Normal Form

A λ -term in head normal form is generally in the form:

$$\lambda x_1 \dots x_n. x M_1 \dots M_m \quad n, m \geq 0$$

In this case x is called the head variable. If $M \equiv \lambda x_1 \dots x_n. (\lambda x. M_0) M_1 \dots M_m$ where $n \geq 0$, $m \geq 1$ then the subterm $(\lambda x. M_0) M_1$ is called the head redex of M . Following are some examples of λ -terms in head normal form:

$$(1) xM \quad (2) \lambda x. x \quad (3) \lambda xy. x((\lambda z. z)y)$$

Normal order: Notice that, an expression in head normal form may contain redexes in argument positions whereas a normal form may not.

1.5 Reduction strategies

Recall that a term is said to be in β -normal form if it has no β -redexes, that is, subterms of the shape $(\lambda x. M)N$. A term has a β -normal form if it can be reduced to a term in β in β -normal form. It is clear that if a term has a β -normal form, then we can get to the β -normal form by exhaustively reducing all β -redexes of the term, then reducing all β -redexes of all resulting terms, and so forth until we get the β -normal form. A **β -reduction strategy** is a function that selects, whenever a term has multiple β -redexes, which one should be reduced first. If a term is in β -normal form, then no β -reduction is to be done. Notice that, a β -reduction strategy may or may not have the property that adhering to the strategy will ensure that we (eventually) reach a β -normal form, if one exists.

In general, there are several different β -reductions possible for a given term. In this project, there are 5 different reduction strategies studied and implemented: normal order, call-by-name, call-by-value, head reduction, and applicative order. Figure X generally summarizes and compares the property of different strategies:

Reduction Strategy	Reduction Order	Reach β -normal form	Form Reached
Normal Order	leftmost outermost	Yes	normal form
Call-by-Name	leftmost outermost not inside an abstraction	No	weak head normal form
Call-by-Value	leftmost innermost not inside an abstraction	No	weak normal form
Head Reduction	reduction inside lambda abstractions but only in head position	No	head normal form
Applicative Order	leftmost innermost	Yes	normal form

Lambda terms are modeled as Haskell constructed data, representing variable names by character:

```
type Name = Char
```

```
data Term = Var Name | Abs Name Term | App Term Term
  deriving (Show, Eq)
```

Operational Semantics

A sequence of computational steps of a valid program is the operational semantics for the programming language. The final step of a terminating sequence returns the resulting value of the program. Operational semantics can be classified into two types: **structural operational semantics (SOS or small-step semantics)** and **natural semantics (NS or big-step semantics)**. The structural operational semantics describes each single step of a computation in a program, while the natural semantics describes the overall resulting value generated by the program. Following, we use both SOS and NS to define the behaviour of each reduction strategy. The small-step semantics gives reduction rules of the strategy, when the big-step semantics describes how the final term is reached.

Recursion

Although a reduction strategy may or may not have the property that ensures it reaches a β -normal form, it would recursively contract the selected redex until no redex (decided by specific reduction strategy) can be contracted. All the following Haskell functions defined corresponding to the reduction strategy model the small-step semantics. Therefore, in order to recursively reduce a λ -term, an auxiliary function `recur_eval :: Term -> IO()` is defined:

```
recur_eval :: Term -> IO()
recur_eval t | evalcbn t == t = putStrLn("")
              | otherwise = do
                  putStrLn("==> "++ toString (evalcbn t))
                  recur_eval (evalcbn t)
```

The Haskell function above defines the recursive reduction using Call-by-Name strategy. It uses the small-step function defined in Section 1.5.2. The recursion will terminate if the input λ -term cannot be reduced further, otherwise, it would print-out the term after a single step reduction and continue reducing the resulting term.

1.5.1 Normal Order Reduction

The normal order reduction $e \xrightarrow{no} e'$ continually applies the rule for β -reduction on the redex in leftmost outermost position until no more β -reduction can be performed. At that point, the resulting term is in normal form. When reducing an application $(e_1 e_2)$, the function term e_1 must be reduced using call-by-name. Since the strategy is outermost, when e_1 is reduced to an abstraction $(\lambda x.e)$, then the redex $((\lambda x.e) e_2)$ must be reduced before redexes in e .

$$\begin{array}{c}
 \frac{\llbracket x \rrbracket_{no}^{ls} = x}{\llbracket e \rrbracket_{no}^{ls} = e'} \\
 \frac{\llbracket e \rrbracket_{no}^{ls} = e'}{\llbracket (\lambda x.e) \rrbracket_{no}^{ls} = (\lambda x.e')} \\
 \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = (\lambda x.e) \quad \llbracket e[e_2/x] \rrbracket_{no}^{ls} = e'}{\llbracket (e_1 e_2) \rrbracket_{no}^{ls} = e'} \\
 \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = e'_1 \neq (\lambda x.e) \quad \llbracket e'_1 \rrbracket_{no}^{ls} = e''_1 \quad \llbracket e_2 \rrbracket_{no}^{ls} = e'_2}{\llbracket (e_1 e_2) \rrbracket_{no}^{ls} = (e''_1 e'_2)}
 \end{array}$$

Normal Order: Big-step operational semantics

$$\frac{}{(\lambda x.M)N \rightarrow_{\beta} M[N/x]} \quad \frac{M \rightarrow_{\beta} N}{MP \rightarrow_{\beta} NP} \quad \frac{M \rightarrow_{\beta} N}{PM \rightarrow_{\beta} PN} (P \text{ contains no redex}) \quad \frac{M \rightarrow_{\beta} N}{(\lambda x.M) \rightarrow_{\beta} (\lambda x.N)}$$

Normal Order: Small-step operational semantics

Normal order reduction is *normalizing*, since it reduces the λ -term until there is no redex, redex in abstractions is also contracted. So, the normal order reduction will terminate with the normal form as result.

The corresponding Haskell function `evalnormal :: Term -> Term` below implements the normal order reduction. Notice that, it uses the function `evalcbn` defined in 1.5.2:

```
--normal order reduction
evalnormal :: Term -> Term
evalnormal (Var x) = Var x
evalnormal (Abs x t) = (Abs x (evalnormal t))
evalnormal (App (Abs x t) v) = subs t (Var x, v)
evalnormal (App x t) | x == evalnormal x = (App x (evalnormal t))
                    | otherwise = (App (evalnormal x) t)
```

The first function clause above handles variables x . It is not a β -reduction step, since a variable is not a redex. It returns itself which indicates no reduction can be done. It also conforms the pattern matching mechanism in Haskell. Similarly, the second function clause handles abstractions $(\lambda x.e)$ and implements the fourth small-step rule. The third function handles applications $(e_1 e_2)$, when e_1 is an abstraction, then the substitution will take place. It applies argument to function and returns the resulting value. Finally, the fourth function handles applications $(e_1 e_2)$ when e_1 is not an abstraction. Since the normal reduction is leftmost outermost, it reduces the function e_1 first. If e_1 cannot be reduced (returns itself), then it operates the argument e_2 .

Example 1.5.1 Following is the running example of normal order reduction in Haskell. We use “/” to represent the λ symbol:

```
(/a.a)(/b.b)((/x.x)(/y.(/z.z)w))
==> (/b.b)((/x.x)(/y.(/z.z)w))
==> (/x.x)(/y.(/z.z)w)
==> /y.(/z.z)w
==> /y.w
Reduced to:/y.w
```

As we can see, it uses leftmost outermost strategy. The leftmost outermost redex $(\lambda a.a)(\lambda b.b)$ is first reduced. Then the whole reduced term is a redex $b[(\lambda x.x)(\lambda y.(\lambda z.z)w)/b]$, the argument $((\lambda x.x)(\lambda y.(\lambda z.z)w))$ replaces the free occurrence of b in b . After that, the same substitution takes place. Finally, it reaches into the redex $(\lambda z.z)w$ in the abstraction $\lambda y.(\lambda z.z)w$ and generates the resulting term $\lambda y.w$ in normal form.

1.5.2 Call-by-Name Reduction

The Call-by-Name reduction $e \xrightarrow{cbn} e'$ reduces the leftmost outermost redex not inside a lambda abstraction first. That is, the arguments to a function are not evaluated before the function is called. The redex in e of the abstraction $(\lambda x.e)$ will never be reduced.

$$\begin{aligned} \llbracket x \rrbracket_{cbn}^{ls} &= x \\ \llbracket (\lambda x.e) \rrbracket_{cbn}^{ls} &= (\lambda x.e) \\ \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = (\lambda x.e) \quad \llbracket e[e_2/x] \rrbracket_{cbn}^{ls} = e'}{\llbracket (e_1 e_2) \rrbracket_{cbn}^{ls} = e'} \\ \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = e'_1 \neq (\lambda x.e)}{\llbracket (e_1 e_2) \rrbracket_{cbn}^{ls} = (e'_1 e_2)} \end{aligned}$$

Call-by-Name: Big-step operational semantics

$$\frac{}{(\lambda x.M)N \rightarrow_{\beta} M[N/x]} \quad \frac{M \rightarrow_{\beta} N}{MP \rightarrow_{\beta} NP}$$

Call-by-Name: Small-step operational semantics

The Call-by-Name reduction generates term in weak head normal form. A lambda expression is in weak head normal form if it is a head normal form or any lambda abstraction.

The corresponding Haskell function `evalcbn :: Term -> Term` below implements the Call-by-Name reduction:

```
--call-by-name reduction
evalcbn :: Term -> Term
evalcbn (Var x) = Var x
evalcbn (Abs x t) = (Abs x t)
evalcbn (App (Abs x t) y) = subs t (Var x, y)
evalcbn (App x y) | x == evalcbn x = (App x y)
                  | otherwise = (App (evalcbn x) y)
```

The first two function clauses handle the variables and abstractions, it returns itself since they cannot be reduced under Call-by-Name reduction. The third function implements the first SOS rule which performs substitution. The last function implements the second SOS rule: the argument of an application will not be reduced before the function is called.

Example 1.5.2 The same example in 1.5.1 run in Haskell by the Call-by-Name strategy is as follows:

```
(/a.a)(/b.b)((/x.x)(/y.(/z.z)w))
==> (/b.b)((/x.x)(/y.(/z.z)w))
==> (/x.x)(/y.(/z.z)w)
==> /y.(/z.z)w
Reduced to:/y.(/z.z)w
```

It is easy to see that the first three reduction steps are the same as normal order reduction. In this example, it stops at $\lambda y.(\lambda z.z)w$. Since the difference between normal order and call-by-name is the redex inside abstractions will never be reduced in call-by-name, and the argument will not be reduced before the function is called.

1.5.3 Call-by-Value Reduction

Call-by-Value reduction is the most common reduction strategy. In Call-by-Value, the argument expression is evaluated, and the resulting value is bound to the corresponding variable in the function. It first reduces the leftmost innermost redex not inside an abstraction. It never reduces a redex when the argument is not a *value*. It differs from Call-by-Name only by reducing the argument e_2 of the application $(e_1 e_2)$ before the substitution take place.

$$\begin{aligned} \llbracket x \rrbracket_{cbv}^{ls} &= x \\ \llbracket (\lambda x.e) \rrbracket_{cbv}^{ls} &= (\lambda x.e) \\ \frac{\llbracket e_1 \rrbracket_{cbv}^{ls} = (\lambda x.e) \quad \llbracket e_2 \rrbracket_{cbv}^{ls} = e'_2 \quad \llbracket e[e'_2/x] \rrbracket_{cbv}^{ls} = e'}{\llbracket (e_1 e_2) \rrbracket_{cbv}^{ls} = e'} \\ \frac{\llbracket e_1 \rrbracket_{cbv}^{ls} = e'_1 \neq (\lambda x.e) \quad \llbracket e_2 \rrbracket_{cbv}^{ls} = e'_2}{\llbracket (e_1 e_2) \rrbracket_{cbv}^{ls} = (e'_1 e'_2)} \end{aligned}$$

Call-by-Value: Big-step operational semantics

$$\frac{}{(\lambda x.M)N \rightarrow_{\beta} M[N/x]} (N \text{ in normal form}) \quad \frac{M \rightarrow_{\beta} N}{MP \rightarrow_{\beta} NP} (M \text{ not an abstraction}) \quad \frac{M \rightarrow_{\beta} N}{VM \rightarrow_{\beta} VN} (M \text{ not an abstraction})$$

Call-by-Value: Small-step operational semantics

Notice that, the third rule specifies that the call-by-value reduction always reduces the argument first before the substitution take place. If V is an abstraction, so it cannot be reduced by call-by-name reduction, then it applies to the third rule. If V is not an abstraction, if it can be reduced then it applies to the second rule; if it cannot be reduced, then applies to the third rule.

Call-by-Value generates terms in weak head normal form only. The implementation of the rules by an Haskell function `evalcbv :: Term -> Term` is as follows:

```
--call-by-value reduction
evalcbv :: Term -> Term
evalcbv (Var x) = Var x
evalcbv (Abs x t) = (Abs x t)
evalcbv (App (Abs x t) y) | y == evalcbv y = subs t (Var x, y)
                        | otherwise = App (Abs x t) (evalcbv y)
evalcbv (App x y) | x == evalcbv x = (App x (evalcbv y))
                  | otherwise = (App (evalcbv x) y)
```

The functions are similar to Call-by-Name, the only difference is the argument is reduced before the substitution take place as indicated in the function clause 3.

Example 1.5.3 The sample example run in Haskell by the Call-by-Value strategy is as follows:

```
(/a.a) (/b.b) ((/x.x) (/y. (/z.z)w))
==> (/b.b) ((/x.x) (/y. (/z.z)w))
==> (/b.b) (/y. (/z.z)w)
==> /y. (/z.z)w
Reduced to: /y. (/z.z)w
```

As we can see, the reduction procedure is different from previous. It differs from previous two strategies at the second reduction step. Since the call-by-value reduction always reduces the argument before the substitution takes place, it reduces the redex in the argument $((\lambda x.x)(\lambda y.(\lambda z.z)w))$ before it replaces the free occurrence of b in b .

1.5.4 Head Reduction

The head reduction performs reductions inside lambda abstractions, but only in head position. Notice that, the *head reduction* strategy introduced in this project is the same as defined by Barendregt [1], however it differs from Sessoft's [12] *head spine reduction* which is implemented by Paulson's **headNF** function[9]. In the leftmost head reduction, only head redexes are reduced. A redex $((\lambda x.e_0)e_1)$ is a *head redex* if it is proceeded to the left only by lambda abstractions of non-redexes, as in $\lambda x_1 \dots \lambda x_n. (\lambda x.e_0)e_1 \dots e_m$ when $n \geq 0$ and $m \geq 1$.

$$\begin{aligned} & \frac{}{\llbracket x \rrbracket_{hr}^{ls} = x} \\ & \frac{}{\llbracket e \rrbracket_{hr}^{ls} = e'} \\ & \frac{}{\llbracket (\lambda x.e) \rrbracket_{hr}^{ls} = (\lambda x.e')} \\ & \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = (\lambda x.e) \quad \llbracket e[e_2/x] \rrbracket_{hr}^{ls} = e'}{\llbracket (e_1 e_2) \rrbracket_{hr}^{ls} = e'} \\ & \frac{\llbracket e_1 \rrbracket_{cbn}^{ls} = e'_1 \neq (\lambda x.e)}{\llbracket (e_1 e_2) \rrbracket_{hr}^{ls} = (e'_1 e_2)} \end{aligned}$$

Head Reduction: Big-step operational semantics

$$\frac{}{(\lambda x.M)N \rightarrow_\beta M[N/x]} \quad \frac{M \rightarrow_\beta N}{MP \rightarrow_\beta NP} \quad \frac{M \rightarrow_\beta N}{(\lambda x.M) \rightarrow_\beta (\lambda x.N)}$$

Head Reduction: Small-step operational semantics

The head reduction strategy generates terms in head normal form. Recall that, a term is in head normal form if it has the form: $\lambda x_1 \dots x_n. x M_1 \dots M_m \quad n, m \geq 0$

```
--head reduction
headreduction :: Term -> Term
headreduction (Var x) = Var x
headreduction (Abs x t) = (Abs x (headreduction t))
headreduction (App (Abs x t) y) = subs t (Var x, y)
headreduction (App x y) | x == headreduction x = (App x y)
                        | otherwise = (App (headreduction x) y)
```

The functions are similar to call-by-name reduction, the only difference is the redex inside abstraction is also reduced.

Example 1.5.4 The sample example run by Head Reduction is as follows:

```
(/a.a)(/b.b)((/x.x)(/y.(/z.z)w))
==> (/b.b)((/x.x)(/y.(/z.z)w))
==> (/x.x)(/y.(/z.z)w)
==> /y.(/z.z)w
==> /y.w
Reduced to:/y.w
```

The reduction procedure is the same as normal order reduction. Notice that, in the third reduction step, the redex $(\lambda z.z)w$ is a head redex in the abstraction $\lambda y.(\lambda z.z)w$.

1.5.5 Applicative Order Reduction

Applicative order reduction $e \xrightarrow{ao} e'$ reduces the leftmost innermost redex first. A function's arguments are always reduced before the function is called. Applicative order always attempts to apply functions to normal forms. It differs from Call-by-Value only by reducing also under abstractions:

$$\frac{\begin{array}{c} \llbracket x \rrbracket_{ao}^{ls} = x \\ \llbracket e \rrbracket_{ao}^{ls} = e' \end{array}}{\llbracket (\lambda x.e) \rrbracket_{ao}^{ls} = (\lambda x.e')}$$

$$\frac{\begin{array}{c} \llbracket e_1 \rrbracket_{ao}^{ls} = (\lambda x.e) \quad \llbracket e_2 \rrbracket_{ao}^{ls} = e'_2 \quad \llbracket e[e'_2/x] \rrbracket_{ao}^{ls} = e' \end{array}}{\llbracket (e_1 e_2) \rrbracket_{ao}^{ls} = e'}$$

$$\frac{\begin{array}{c} \llbracket e_1 \rrbracket_{ao}^{ls} = e'_1 \neq (\lambda x.e) \quad \llbracket e_2 \rrbracket_{ao}^{ls} = e'_2 \end{array}}{\llbracket (e_1 e_2) \rrbracket_{ao}^{ls} = (e'_1 e'_2)}$$

Applicative Order: Big-step operational semantics

$$\frac{}{(\lambda x.M)N \rightarrow_\beta M[N/x]} (M, N \text{ in normal form}) \quad \frac{M \rightarrow_\beta N}{MP \rightarrow_\beta NP} \quad \frac{M \rightarrow_\beta N}{PM \rightarrow_\beta PN} (P \text{ contains no redex})$$

$$\frac{M \rightarrow_\beta N}{(\lambda x.M) \rightarrow_\beta (\lambda x.N)}$$

Applicative Order: Small-step operational semantics

Applicative order reduction always generates terms in normal form. Since it also reduce the redexes in abstractions. The applicative order reduction is not normalizing, functions applied to non-normalizing arguments are non-normalizing.

```
--applicative order reduction
apporder :: Term -> Term
apporder (Var x) = Var x
apporder (Abs x t) = (Abs x (apporder t))
apporder (App (Abs x t) y) | apporder (Abs x t) /= (Abs x t)
    = (App (apporder (Abs x t)) y)
    | apporder (Abs x t) == (Abs x t) && apporder y /= y
    = (App (Abs x t) (apporder y))
    | apporder (Abs x t) == (Abs x t) && apporder y == y
    = subs t (Var x, y)
apporder (App x y) | x == apporder x = (App x (apporder y))
    | otherwise = (App (apporder x) y)
```

The function definitions of applicative order reduction is quite different with others. Since it reduces the leftmost **innermost** redex first, and the arguments are always reduced before the function, it needs more ‘guards’ for those cases. As we can see from the third function clause, it uses guards to list 3 possible cases: 1. reduce the function if possible 2. reduce the argument if the function cannot be reduced. 3. both function and argument cannot be reduced, substitution takes place.

Example 1.5.5 The sample example run by Applicative Order reduction in Haskell is as follows:

```
(/a.a)(/b.b)((/x.x)(/y.(/z.z)w))
==> (/b.b)((/x.x)(/y.(/z.z)w))
==> (/b.b)((/x.x)(/y.w))
==> (/b.b)(/y.w)
==> /y.w
Reduced to:/y.w
```

Since applicative order reduction reduces the leftmost innermost redex first, it firstly reduces the redex $(/a.a)(/b.b)$. It always reduces the arguments before the function, so it reduces the argument $((/x.x)(/y.(/z.z)w))$ next until it contains no redex(in normal form). When the argument is in the normal form $(/y.w)$, the substitution takes place and finally it reaches the normalized term $(/y.w)$.

1.6 Comparisons between Different Reduction Strategies

1.7 λ -calculus with Explicit Substitution and Garbage Collection

We currently have introduced the untyped λ -calculus with implicit substitution. When we perform β -reduction on an application $(\lambda x.xz)y$ whose function is an abstraction, the substitution is done in 1 single step: that is, all unbound occurrences of variable x are substituted by the argument y . However, as a term becomes complex, it gives less details about how the substitution is done. In this section, a new calculi with explicit substitution built based on ordinary λ -calculus is introduced. It generates substitutions and treats them as a λxgc -term. A set of explicit substitution rules are defined to reduce a term into normal form.

1.7.1 The explicit substitution calculi

In Bloo et al. [2], a new λ -calculus with explicit substitution and garbage collection is introduced and studied. It retains the use of traditional variable names, specifying terms modulo renaming and includes reduction rules for *explicit garbage collection*. It is a conservative extension of the ordinary λ -calculus from several properties. The λxgc -calculus with syntactic reductions for explicit substitution and garbage collection is represented.

The set of λxgc -terms Λx , ranged over by $MNPQ$, is formally defined by:

Definition 1.7 (λxgc -terms). λxgc -terms are the the class satisfying the following grammar:

$$M ::= x \mid \lambda x.M \mid MN \mid M\langle x := N \rangle$$

where the letters $xyzvw$ range over an infinite set of variables. The λxgc -calculus follows the parentheses omission of original λ -calculus and write $\lambda xy.M$ for $\lambda x.(\lambda y.M)$ and $MN(PQ)$ for $(MN)(PQ)$; explicit substitution is given highest precedence so $\lambda x.MNP\langle x := Q \rangle$ is $\lambda x.((MN)(P\langle x := N \rangle))$.

The λxgc -terms include as a subset the ordinary λ -terms Λ ; a λxgc -terms is ‘pure’ if it is also a λ -term, *i.e.*, if it has no subterms of the form $M\langle x := N \rangle$.

Definition 1.8 (free and bound variables of λxgc -terms) The set of free and bound variables are defined inductively by the following function:

$$\begin{array}{llll} FV_{\text{xgc}}(x) & = & \{x\} & BV_{\text{xgc}}(x) & = & \emptyset \\ FV_{\text{xgc}}(\lambda x.M) & = & FV_{\text{xgc}}(M) \setminus \{x\} & BV_{\text{xgc}}(\lambda x.M) & = & BV(M) \cup \{x\} \\ FV_{\text{xgc}}(MN) & = & FV_{\text{xgc}}(M) \cup FV_{\text{xgc}}(N) & BV_{\text{xgc}}(MN) & = & BV_{\text{xgc}}(M) \cup BV_{\text{xgc}}(N) \\ FV_{\text{xgc}}(M\langle x := N \rangle) & = & (FV_{\text{xgc}}(M) \setminus \{x\}) \cup FV_{\text{xgc}}(N) & BV_{\text{xgc}}(M\langle x := N \rangle) & = & BV_{\text{xgc}}(M) \cup x \\ & & & & & \cup BV_{\text{xgc}}(N) \end{array}$$

The last free and bound variable set definition of substitution is new. All the free occurrences of x in M is actually bounded by x in the substitution $\langle x := N \rangle$. Therefore, the free variables of term M should exclude x .

Definition 1.9 (λxgc -term concepts).

- (**α -equivalent**) Two terms are α -equivalent, written $M \equiv N$, is as for λ -calculus plus (the new) $M\langle x := N \rangle \equiv P\langle y := Q \rangle$ if $N \equiv Q$ and $M[x := z] \equiv P[y := z]$ for $z \notin FV(M)$
- (**Garbage**) The substitution $\langle x := N \rangle$ in $M\langle x := N \rangle$ is called *garbage* if $x \notin FV(M)$

Definition 1.10 (‘raw’ reduction step of xgc)

- **Substitution generation**, \xrightarrow{b} , is

$$(\lambda x.M)N \rightarrow M\langle x := N \rangle \quad (\text{b})$$

- **Explicit Substitution**, \xrightarrow{x} is

$$\begin{array}{ll} x\langle x := N \rangle \rightarrow N & (\text{xv}) \\ x\langle y := N \rangle \rightarrow x & \text{if } x \neq y \quad (\text{xvgc}) \\ (\lambda x.M)\langle y := N \rangle \rightarrow \lambda z.M[x := z]\langle y := N \rangle & \text{if } x \in FV(N) \ \& \ x \neq y, z \text{ new} \quad (\text{xaba}) \\ (\lambda x.M)\langle y := N \rangle \rightarrow \lambda x.M\langle y := N \rangle & \text{if } x \notin FV(N) \ \& \ x \neq y \quad (\text{xab}) \\ (M_1M_2)\langle y := N \rangle \rightarrow M_2\langle y := N \rangle M_2\langle y := N \rangle & (\text{xap}) \end{array}$$

The **xaba** rule defines the α -conversion in λxgc -calculus. For the term $(\lambda y.xy)\langle x := y \rangle$, y is bound in $\lambda y.xy$ and free in $\langle x := y \rangle$. If x is replaced by y in $\lambda y.xy$, it becomes $\lambda y.yy$ and *variable capture* occurs. In this case, the bound variable should be renamed.

- **Garbage Collection**, \vec{gc} , is

$$M\langle x := N \rangle \rightarrow M \quad \text{if } x \notin FV(M) \quad (gc)$$

- \xrightarrow{xgc} is the union of \xrightarrow{x} and \xrightarrow{gc} .
- \xrightarrow{bxgc} is the union of \xrightarrow{b} and \xrightarrow{xgc} .

1.7.2 Reduction in λxgc -calculus

Recall that if a λxgc -term does not contain any subterms of the form $M\langle x := N \rangle$, then it is ‘pure’. The abstraction in λxgc -calculus is different from the ordinary λ -calculus: an abstraction in λxgc -calculus can also in the form $\lambda x.M\langle x := N \rangle$. In this case, β -reduction cannot be applied to this term since it can only apply to λ -terms. Therefore, it should perform explicit substitution steps until it is a ‘pure’ λxgc -term.

Example 1.7.1 The sample run of normal order reduction with explicit substitution and garbage collection is as following(‘ \backslash ’ as λ):

(1)	$(\backslash x. (\backslash y. x)x) (\backslash z. q)$	whole term is a redex
(2)	$\Rightarrow ((\backslash y. x)x) \langle x := \backslash z. q \rangle$	xgc-reduction
(3)	$\Rightarrow ((\backslash y. x) \langle x := \backslash z. q \rangle) (x \langle x := \backslash z. q \rangle)$	
(4)	$\Rightarrow (\backslash y. x \langle x := \backslash z. q \rangle) (x \langle x := \backslash z. q \rangle)$	
(5)	$\Rightarrow (x \langle x := \backslash z. q \rangle) \langle y := x \langle x := \backslash z. q \rangle \rangle$	garbage collection
(6)	$\Rightarrow x \langle x := \backslash z. q \rangle$	
(7)	$\Rightarrow \backslash z. q$	

Recall that, the normal order reduction reduces the leftmost outermost redex first. At the beginning, the whole term $(\backslash x. (\backslash y. x)x) (\backslash z. q)$ is a redex, and the function of the application is an abstraction. Therefore, the *substitution generation* \vec{b} is performed. Further, the substitution is pushed into the application and multiple explicit substitution steps are performed. At step (4), another *substitution generation* \vec{b} is performed. Since there is no y in $x \langle x := \backslash z. q \rangle$, the substitution $\langle y := x \langle x := \backslash z. q \rangle \rangle$ is the garbage and removed. Finally, x is substituted by $\backslash z. q$.

Example 1.7.2 Below is another sample run by applicative order reduction with explicit substitution and garbage collection:

(1)	$(\backslash x. (\backslash y. x)x) (\backslash z. q)$	reduce the innermost redex $(\backslash y. x)x$
(2)	$\Rightarrow (\backslash x. x \langle y := x \rangle) (\backslash z. q)$	garbage collection
(3)	$\Rightarrow (\backslash x. x) (\backslash z. q)$	
(4)	$\Rightarrow x \langle x := \backslash z. q \rangle$	
(5)	$\Rightarrow \backslash z. q$	

Recall that, the applicative order reduction reduces the leftmost innermost redex first. The redex $(\backslash y. x)x$ is the leftmost and innermost redex at beginning. Therefore, the *substitution generation* is performed inside the function. Since there is no free occurrence of y in x , the substitution $\langle y := x \rangle$ is the garbage. At step (3), another *substitution generation* is performed and finally we get the term $\backslash z. q$.

1.8 Normal β -reduction vs Explicit Substitution and garbage collection

As might be expected, it takes more reduction steps to reduce a λ -term to normal form using \xrightarrow{bxgc} than it takes using $\xrightarrow{\beta}$. In this section, it compares the difference between the number of

reduction steps needed by normal β -reduction and the explicit substitution and garbage collection. In addition, it is separated into two parts: β -reduction versus explicit substitution (to see the number of steps increased), and explicit substitution versus \xrightarrow{bxgc} (to demonstrate the efficiency improved by garbage collection).

1.8.1 $\xrightarrow{\beta}$ vs \xrightarrow{bx}

To illustrate the differences, we use normal order reduction and list the number of steps required to normalize the term by normal reduction and explicit substitution. There is also a line chart to demonstrate the gap between these two methods. We firstly use some simple terms that cost several steps and then we use the term in the form $\lambda v.(\lambda x.I(Ix))(Iv)$; we gradually enlarge the term I , then the reduction steps needed grows.

λ -term	Normal Reduction	Explicit Substitution
$(\lambda x.x)y$	1	2
$(\lambda abc.ac(bc))(\lambda xy.x)$	3	12
$I = \lambda y.y$	4	14
$(\lambda fx.f(fx))(\lambda fx.f(fx))$	6	45
$I = (\lambda yw.w)z$	7	27
$I = (\lambda ywv.wv)z$	9	67
$I = (\lambda ywvs.wvs)z$	11	123
$I = (\lambda ywvsd.wvsd)z$	13	195

Table 1.1: Reduction steps needed for normal reduction and explicit substitution

In Table 1.1, it lists 8 terms from simple to complex. The simple application $(\lambda x.x)y$ only takes one reduction step that substitutes the free occurrence of x in $x[y/x]$ by y . The explicit substitution takes two steps: substitution generation and substitution. The second term $(\lambda abc.ac(bc))(\lambda xy.x)$ needs more steps for explicit substitution since it has to push the substitution into abstraction and further into application until it reaches a variable. The sample run of this term with explicit substitution can also be found in Figure 5.9. When the term $(\lambda fx.f(fx))(\lambda fx.f(fx))$ is normalized with explicit substitution, it needs 45 steps whereas all the β -reduction paths of it only consists 6 steps. Finally, a series of terms in the form $\lambda v.(\lambda x.I(Ix))(Iv)$ are used. We gradually enlarge the term I to increase the number of β -reduction steps and to evaluate the gap between two calculus. As it illustrated in the table, the number of steps needed for explicit substitution grows larger than the square of normal β -reduction when $I = (\lambda ywvs.wvs)z$. When $I = (\lambda ywvsd.wvsd)z$, the number of steps by explicit substitution is 195 which is much larger than the square of the number of β -reductions 13.

We can see that, the number of steps needed by explicit substitution mainly depends on the number of β -reduction required and the structure of abstraction's body. It needs more steps to push a substitution into an application if the structure is complex. For example, the λxgc -term $(xyz)\langle x := z \rangle$ needs 2 steps to push the substitution into the application xyz : $(xy)\langle x := z \rangle z \langle x := z \rangle$ and then $x\langle x := z \rangle y \langle x := z \rangle z \langle x := z \rangle$. If an application only contains n variables, then it needs $n - 1$ steps to push a substitution into it, and it also takes n steps to reduce each λxgc -term in form $x\langle x := y \rangle$. It needs more steps if the application also contains abstractions with complex body.

The figure 1.1 illustrates the difference in the number of steps needed between normal β -reduction and \xrightarrow{bx} . As the term I grows larger, the number of steps in normal β -reduction grows linearly, while xgc grows approximately exponentially. We can see that the β -reduction steps needed for term $(\lambda fx.f(fx))(\lambda fx.f(fx))$ is less than the steps needed when $I = (\lambda yw.w)z$, however, the number of explicit substitution steps is larger. The number of reduction steps mainly depends on the free occurrences of bound variables in an abstraction. For the term $(\lambda fx.f(fx))$, the free occurrence of f and x in $f(fx)$ lead to repeated substitution and increase the number of reduction steps.

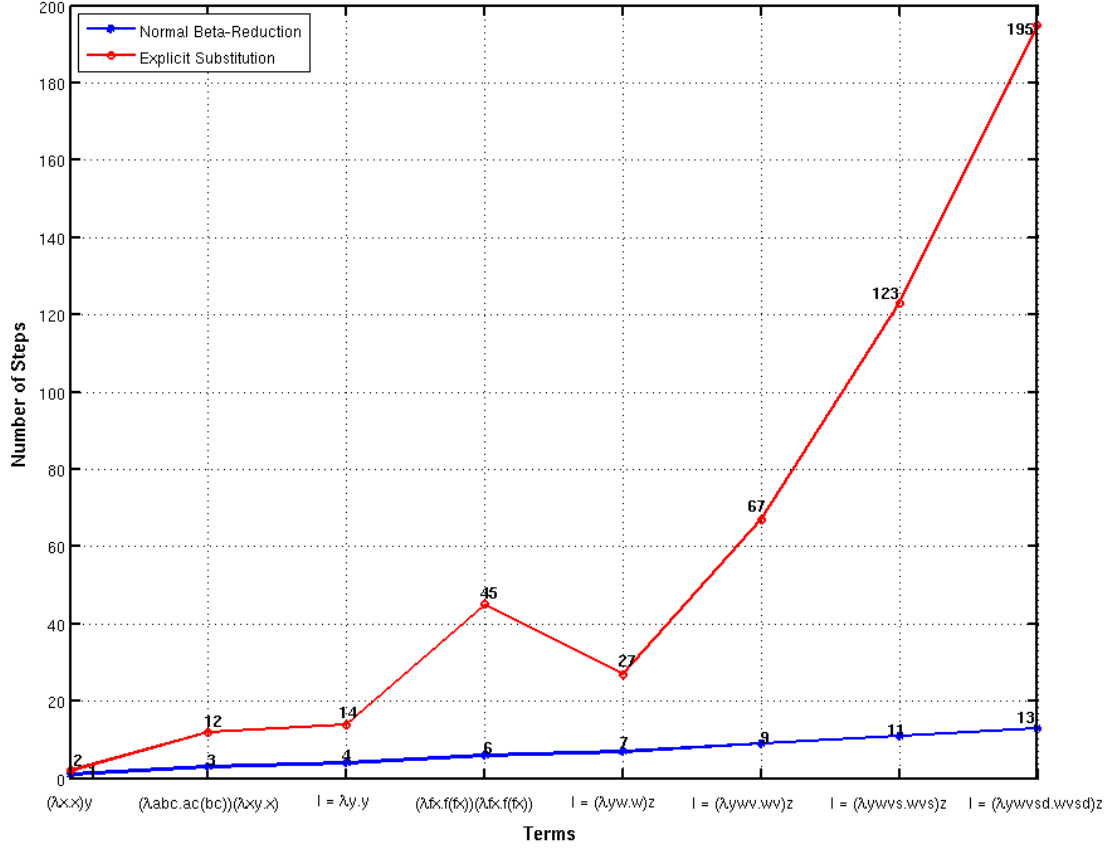


Figure 1.1: Reduction steps needed for normal β -reduction and explicit substitution

1.8.2 $\xrightarrow{\beta}$ vs \xrightarrow{bx} vs \xrightarrow{bxgc}

The explicit garbage collection simplifies several proofs. For the λxgc -term $(xy)\langle z := M \rangle$, since there is no free occurrence of z in the term xy , so the substitution $\langle z := M \rangle$ can be removed. Then it no longer needs to be pushed into the application xy which costs more explicit substitution steps. The following table and chart would show the number of steps decreased, or in other words, the efficiency improved by garbage collection.

λ -term	Normal Reduction	Explicit Substitution	$bxgc$
$(\lambda x.x)y$	1	2	2
$(\lambda abc.ac(bc))(\lambda xy.x)$	3	12	12
$I = \lambda y.y$	4	14	12
$(\lambda fx.f(fx))(\lambda fx.f(fx))$	6	45	45
$I = (\lambda yw.w)z$	7	27	18
$I = (\lambda ywv.wv)z$	9	67	35
$I = (\lambda ywvs.wvs)z$	11	123	58
$I = (\lambda ywvsd.wvsd)z$	13	195	87

Table 1.2: Reduction steps needed for normal reduction, explicit substitution, and $bxgc$

Compared with Table 1.1, Table 1.2 lists additional the number of steps needed by explicit substitution and garbage collection. As we can see, there is no garbage collection performed for the term $(\lambda x.x)y$ and $(\lambda abc.ac(bc))(\lambda xy.x)$. When $I = \lambda y.y$, the garbage collection takes place at the 2nd and 6th reduction step when the term is $\backslash v.((\backslash y.y)\langle x := (\backslash y.y)v \rangle)((\backslash y.y)x)\langle x := (\backslash y.y)v \rangle$ and $\backslash v.((\backslash y.y)\langle x := (\backslash y.y)v \rangle)(x\langle x := (\backslash y.y)v \rangle)$. There is no occurrence of x in $\backslash y.y$ in both

cases. So the substitution is the garbage and is removed. Therefore, the substitution would not be pushed into the abstraction and it reduces 2 reduction steps.

For the term $(\lambda f x. f(fx))(\lambda f x. f(fx))$, there is no garbage collection performed because of the repeated occurrence of f and x . When the term I is gradually enlarged, the number of steps reduced by garbage collection becomes larger. It decreases more than a half when $I = (\lambda y w v. wv)z$. The more the number of reduction steps needed by explicit substitution, the more steps reduced by garbage collection. We can see the difference in Figure 1.2

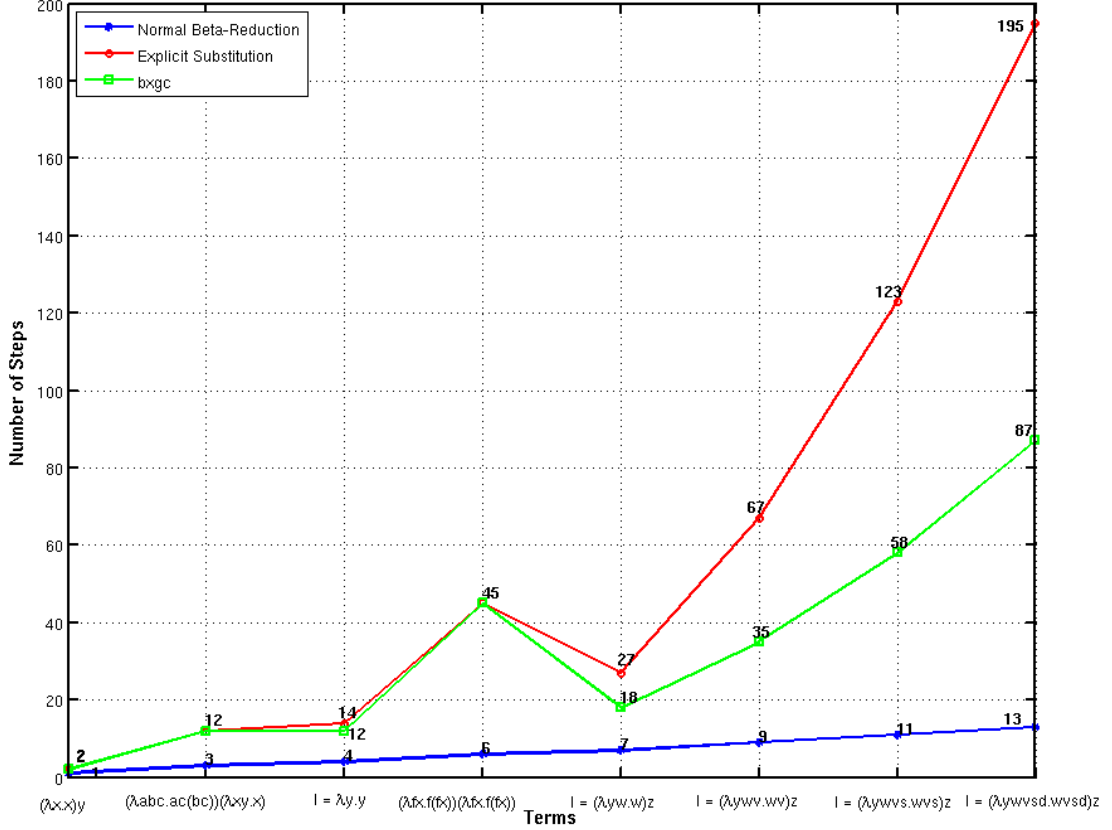


Figure 1.2: Reduction steps needed for normal reduction, bx, and bxgc

The figure 1.2 illustrates the difference in the number of steps needed among normal β -reduction, \xrightarrow{bx} , and \xrightarrow{bxgc} . There is not too much difference between \xrightarrow{bx} and \xrightarrow{bxgc} for the first 4 terms. The garbage collection has a significant impact as the term I grows larger. As we can see from the graph, the difference between \xrightarrow{bx} and \xrightarrow{bxgc} becomes larger from $I = (\lambda yw.w)z$ to $I = (\lambda ywvsd.wvsd)z$.

Although the garbage collection can remove any useless substitutions from λxgc -terms and avoid redundant explicit substitution steps. It takes much more reduction steps than normal β -reduction. The number of steps needed by explicit substitution and garbage collection is approximately $n \log n$ (only rough estimated).

It is worth mentioning that the reduction path of a term by different reduction strategies also affects the number of reduction steps needed for \xrightarrow{bx} and \xrightarrow{bxgc} . The shortest \xrightarrow{bxgc} path of the term $I = \lambda y.y$ is 8, using applicative order reduction. It costs 12 steps for the normal order reduction whereas all the β -reduction paths of it consist of 4 steps.

Chapter 2

The Curry type assignment system

A type system consists a set of rules that assign a property called a *type* to the various constructs — such as variable, expression, functions or modules. There are several reasons to add types to a program, one of the most important reason to add type properties to a program is to reduce bugs. It is possible to build an abstract interpretation of programs by treating terms as objects. The abstractions can be regarded as the interface of objects, by checking whether the interfaces have been connected properly we can decide whether the program is defined in a consistent way. Moreover, type systems provide a form of documentation and improve the readability of a program. Since a program can be abstracted, there are less information in the structure of a program. Furthermore, it also enables independent compilation before the implementation of a program runs. The types of functions and arguments can be checked during the generation of codes. Type-checking during implementation enables dynamic debugging which largely reduces the debugging workload at run-time. If a program is error-free, it is safe to run: "Typed programs cannot go wrong". It also allows multiple dispatch which is the feature of some objected-oriented programming languages in which a function or method can be dynamically dispatched based on the run time type of more than one of its arguments.

An exmaple of a type system is the Java language. A Java program consists of classes which contains a set of function definitions. A function is invoked by an instance of a class that the function belongs to. The interface of a function states the type of the return value, the name of the function and a list of values that are passed to the function. The code of an invoking function states the object instance on which this particular method is to be invoked, and the name of this function along with the names of variables that hold values to pass to it. The Java compiler checks the type of each argument that passed into the function, against the type declared for each variable in the interface of the invoked function. If the types do not match, the compiler throws a compile-time error.

The lambda calculus as treated in Chapter 1 is referred to as a *type-free* theory. Because every expression may be applied to every other expression. For example, the $\lambda x.x$ may be applied to any argument x and generates the same x .

The typed version of lambda calculus(or called Combinatory Logic, a variant of the lambda calculus) is introduced essentially in Curry[4] and in Church[3]. Types are objects of a syntactic nature and may be assigned to lambda terms. If M is a lambda term and a type A is assigned to M , then we say ' M has type A '; the notation used is ' $M : A$ '. For example, in a typed system one has $\lambda x.x : (A \rightarrow A)$, which means the function $\lambda x.x$ should take an argument in type A and return a value of type A . In general, $A \rightarrow B$ is the type of functions from A to B .

2.1 The System $\lambda \rightarrow$ -Curry

In Curry and Feys[6] and Curry et al[5], the type assignment theory was modified in a natural way to the lambda calculus assigning elements of a given set \mathbb{T} of types to type free lambda terms. For this reason these calculi are sometimes called systems of type assignment. If the type $\sigma \in \mathbb{T}$ is assigned to the term $M \in \Lambda$ which writes as $\vdash M : \sigma$, sometime with a under subscript such

as \vdash_λ to denote the particular system. Usually a set of assumptions Γ is needed to derive a type assignment write as $\Gamma \vdash M : \sigma$ (read as ‘ Γ yields M in σ ’).

Definition 2.1 (i) The set of *types*, notation \mathbb{T} ,

Notation. (i) If $\sigma_1, \dots, \sigma_n \in \mathbb{T}$ then

$$\sigma_1 \rightarrow \sigma_2 \rightarrow \dots \rightarrow \sigma_n$$

stands for

$$(\sigma_1 \rightarrow (\sigma_2 \rightarrow \dots \rightarrow (\sigma_{n-1} \rightarrow \sigma_n)))$$

we use association to the right, and the rightmost and outermost parentheses are normally omitted.

(ii) $\alpha, \beta, \gamma, \dots$ denote arbitrary type variables.

Definition 2.2 (i) A *statement* is of the form $M : \sigma$, where $M \in \Lambda$ and $\sigma \in \mathbb{T}$ (pronounced as ‘ M in σ ’). M is called the *subject* and A the *predicate* of $M : A$.

(ii) A *context* Γ is a set of statements with only distinct (term) variables as subjects; In [13], the notion $\Gamma, M : \sigma$ is used for the context $\Gamma \cup \{M : \sigma\}$ where either $M : \sigma \in \Gamma$ or M does not occur in Γ , and we use $M : \sigma$ as shorthand for $\emptyset, M : \sigma$.

The notion of *context* will be used to collect all statements used for the free variables of a term when typing that term[13].

Definition 2.3 (CF. [[4], [6]]) *Curry type assignment* and *derivations* can be defined using following derivation rules.

$$\begin{aligned} (1) : & \frac{}{\Gamma, M : \sigma \vdash_c M : \sigma} \\ (2) : & \frac{\Gamma, x : \sigma \vdash_c M : \beta}{\Gamma \vdash_c (\lambda x. M) : \sigma \rightarrow \beta} \\ (3) : & \frac{\Gamma \vdash_c M_1 : \sigma \rightarrow \beta \quad \Gamma \vdash_c M_2 : \sigma}{\Gamma \vdash_c (M_1 M_2) : \beta} \end{aligned}$$

It is built up from assumptions $x : \sigma$, and it forms a natural deduction system. Following is an example that shows how a statement can be derived by assumptions using deduction rules defined above:

Example 2.1.1 (i) Let $\sigma \in \mathbb{T}$. Then $\Gamma \vdash \lambda f x. f(fx) : (\sigma \rightarrow \sigma) \rightarrow \sigma \rightarrow \sigma$, is shown by the following derivation:

$$\begin{aligned} & \frac{}{\Gamma, f : \sigma \rightarrow \sigma \vdash f : \sigma \rightarrow \sigma} \quad (1) \quad \frac{\frac{}{\Gamma, f : \sigma \rightarrow \sigma \vdash f : \sigma \rightarrow \sigma} \quad \frac{}{\Gamma, x : \sigma \vdash x : \sigma}}{\Gamma \vdash fx : \sigma} \\ & \frac{\Gamma \vdash f(fx) : \sigma}{\Gamma \vdash \lambda x. f(fx) : \sigma \rightarrow \sigma} \quad (1) \\ & \frac{\Gamma \vdash \lambda x. f(fx) : \sigma \rightarrow \sigma}{\Gamma \vdash \lambda f x. f(fx) : (\sigma \rightarrow \sigma) \rightarrow \sigma \rightarrow \sigma} \quad (2) \end{aligned}$$

(ii) Notice that, we cannot type ‘*self-application*’ xx . In order to type the application xx , the derivation should be as following:

$$\frac{\frac{}{\Gamma \vdash x : A \rightarrow B} \quad \frac{}{\Gamma \vdash x : A}}{\Gamma \vdash xx : B}$$

As we can see, the term variable x is assigned to two types: $A \rightarrow B$ and A , and either of them can be substituted by the other one. So, $A \rightarrow B \neq A$, and there are two statements of subject x in context Γ with distinct predicates. Those two statements would be conflict according to the definition of context. Therefore, the ‘self-application’ is untypable.

2.2 Properties of $\lambda \rightarrow$

Lemma 2.1 GENERATION LEMMA

- (i) $\Gamma \vdash x : \sigma \Rightarrow (x : \sigma) \in \Gamma$
- (ii) $\Gamma \vdash MN : \gamma \Rightarrow \exists \sigma [\Gamma \vdash M : (\sigma \rightarrow \gamma) \ \& \ \Gamma \vdash N : \sigma]$
- (iii) $\Gamma \vdash \lambda x.M : \rho \Rightarrow \exists \sigma, \gamma [\Gamma, x : \sigma \vdash M : \gamma \ \& \ \rho \equiv (\sigma \rightarrow \gamma)]$

Lemma 2.2 SUBSTITUTION LEMMA

$$\exists \sigma [\Gamma, x : \sigma \vdash M : \gamma \ \& \ \Gamma \vdash N : \sigma] \Rightarrow M[x := N] : \gamma$$

Theorem 2.1 SUBJECT REDUCTION THEOREM

Suppose $M \rightarrow M'$. Then

$$\Gamma \vdash M : \sigma \Rightarrow \Gamma \vdash M' : \sigma$$

PROOF.

$$\begin{array}{ll} \Gamma \vdash (\lambda x.M)N : \sigma & \Rightarrow (\text{GenerationLemma}) \\ \Gamma \vdash (\lambda x.M) : \gamma \rightarrow \sigma \ \& \ \Gamma \vdash N : \gamma & \Rightarrow (\text{GenerationLemma}) \\ \Gamma, x : \gamma \vdash M : \sigma \ \& \ \Gamma \vdash N : \gamma & \Rightarrow (\text{SubstitutionLemma}) \\ \Gamma \vdash M[x := N] : \sigma & \end{array}$$

2.3 The principal type property

Definition 2.4 (Type substitution (van Bakel[13])) The substitution $(\varphi \mapsto C) :$, where φ is a type variable and $C \in \mathbb{T}$ is defined as following:

$$\begin{array}{ll} (\varphi \mapsto C)\varphi & = C \\ (\varphi \mapsto C)\varphi' & = \varphi', \text{ if } \varphi' \neq \varphi \\ (\varphi \mapsto C)A \rightarrow B & = ((\varphi \mapsto C)A) \rightarrow ((\varphi \mapsto C)B) \end{array}$$

Definition 2.5 Let Id_s be the substitution that replaces all type variables by themselves:

(i) Robinson's unification algorithm. Unification of Curry types is defined by:

$$\begin{array}{lll} \text{unify } \varphi & \varphi & = (\varphi \mapsto \varphi) \\ \text{unify } \varphi & B & = (\varphi \mapsto B), \text{ if } \varphi \text{ does not occur in } B \\ \text{unify } A & \varphi & = \text{unify } \varphi \ A \\ \text{unify } (A \rightarrow B) & (C \rightarrow D) & = S_1 \circ S_2 \\ & & \text{where } S_1 = \text{unify } A \ C \\ & & S_2 = \text{unify } (S_1 B) (S_1 D) \end{array}$$

(ii) UnifyContexts

$$\begin{array}{lll} \text{UnifyContexts } (\Gamma_0, x : A) & (\Gamma_1, x : B) & = S_1 \circ S_2 \\ & & \text{where } S_1 = \text{unify } A \ B \\ & & S_2 = \text{UnifyContexts } (S_1 \Gamma_0) (S_1 \Gamma_1) \\ \text{UnifyContexts } (\Gamma_0, x : A) & \Gamma_1 & = \text{UnifyContexts } \Gamma_0 \ \Gamma_1, \text{ if } x \text{ does not occur in } \Gamma_1 \\ \text{UnifyContexts } \emptyset & \Gamma_1 & = Id_s \end{array}$$

Definition 2.6 Principle Pair Algorithm for Λ

$$\begin{aligned}
pp_c \ x &= \langle x : \varphi, \varphi \rangle \\
&\text{where } \varphi = \text{fresh} \\
pp_c \ (\lambda x.M) &= \langle \Pi, A \rightarrow P \rangle, \text{ if } pp_c M = \langle \Pi \cup \{x : A\}, P \rangle \\
&\quad \langle \Pi, \varphi \rightarrow P \rangle, \text{ if } pp_c M = \langle \Pi, P \rangle \quad x \notin \Pi \\
&\quad \text{where } \varphi = \text{fresh} \\
pp_c \ (MN) &= S_1 \circ S_2 \langle \Pi_1 \cup \Pi_2, \varphi \rangle \\
&\text{where } \varphi = \text{fresh} \\
&\quad \langle \Pi_1, P_1 \rangle = pp_c M \\
&\quad \langle \Pi_2, P_2 \rangle = pp_c N \\
&\quad S_1 = \text{unify } P_1 \ (P_2 \rightarrow \varphi) \\
&\quad S_2 = \text{UnifyContexts } (S_1 \Pi_1) \ (S_1 \Pi_2)
\end{aligned}$$

Definition 2.7 Principle Pair Algorithm for Λx

$$\begin{aligned}
pp_c \ (M \langle x := N \rangle) &= S_1 \circ S_2 \langle \Pi_1 \cup \Pi_2, P_1 \rangle \\
&\text{where } P_3 = \text{fresh} \quad \text{if } x \notin FV(M) \\
&\quad \Pi_1 = \Pi_3 \cup \{x : P_3\} \quad \text{if } x \in FV(M) \\
&\quad \langle \Pi_1, P_1 \rangle = pp_c M \\
&\quad \langle \Pi_2, P_2 \rangle = pp_c N \\
&\quad S_1 = \text{unify } P_2 \ P_3 \\
&\quad S_2 = \text{UnifyContexts } (S_1 \Pi_1) \ (S_1 \Pi_2)
\end{aligned}$$

$$(4) : \frac{\Gamma, x : \sigma \vdash M : \beta \quad \Gamma \vdash_c N : \sigma}{\Gamma \vdash_c M \langle x := N \rangle \beta}$$

Chapter 3

A Purely-Functional Programming Language — Haskell

This chapter will introduce functional programming and a purely-functional programming language-Haskell. It will demonstrate how the lambda calculus and Curry type assignment system is implemented in Haskell by self-defined data types, pattern matching, list processing, and recursive functions. As an implementation and usable software, the issues and weaknesses would also be discussed in this chapter.

3.1 Functional Programming

Functional programming uses functions, other than objects or procedures, as the fundamental building blocks of a program. It is a programming paradigm that treats computation as the evaluation of expressions and avoids mutable data and state, in contrast with imperative programming where programs are composed of statements which change global state when executed. Functional programming has its roots in lambda calculus. Lambda calculus provides a theoretical framework for describing functions and their evaluation. Although it is a mathematical abstraction rather than a programming language, it forms the basis of almost all functional programming languages. There are several functional programming languages such as Miranda, Haskell, and OCaml that is defined based on lambda calculus.

3.2 Haskell

Haskell is an advanced purely-functional programming language, which has non-strict semantics and strong static typing. In imperative languages(C, C++, Java...), a sequence of tasks are generated and executed in a program. When they are executed, it can change state. It has a control flow structures for doing some action several times. In purely functional programming, we do not need to tell the computer what to do but rather to tell it what each function is. It also cannot set a variables to something and then set it to something else later. In purely functional languages, a function has no side-effects. The only task for a function is to take the arguments, perform operations, and return the result.

Haskell is lazy, which means that it will not execute functions and calculate things unless it is specifically told to do so. If you compile and run a haskell program, it does not call any functions nor perform any execution. The only thing you can do is call a specific function(it may call other functions) and pass parameters into it, then that function is executed and a result is returned.

Haskell is statically typed. When we compile a haskell program, the compiler knows which variable is a number, which is a string and so on. That means errors can be found at compile time. Haskell uses a type system that has type inference. That means we do not have to explicitly label every piece of code with a type because the type system can automatically type it. Type inference also allows the code to be more general. If we add two parameters together without explicitly stating their types, the function will work on any two parameters that can be added as numbers.

Haskell is elegant and concise, because it uses a lot of high level concepts. Haskell programs are usually shorter than their imperative language equivalents. In addition, shorter programs are easier to develop than longer ones and have less bugs.

In this project, Haskell is self-studied thanks to the course note of COMP120.1 of Imperial College. In this section, some main features of Haskell introduced in Miran[8], would be introduced with examples.

3.2.1 Pattern Matching

Pattern matching consists of listing patterns to which some data should conform and then checking to see if it does conform to any of them and deconstructing the data according to those patterns. When defining functions, we can define separate function bodies for different patterns. This leads to really neat code that is simple and readable. We can pattern match on any data type. A successful match binds the formal parameters in the pattern. The matching process itself occurs “top-down” and “left-right”. Failure of a pattern anywhere in one function body results in failure of the whole matching process, and the next function body is then tried to match. For instance, if `[2, 3]` is matched against `[1, a]`, then 2 fails to match with 1, so the result is a failed match. Following is another example that defines a function using pattern matching:

```
factorial :: (Integral a) => a -> a
factorial 1 = 1
factorial n = n * factorial (n-1)
```

The code above defines a function that calculate the factorial of the input. It uses pattern matching that lists all the possible patterns that can be matched. Here, only two cases, 1 or n. It firstly matches the first function body to check whether the input is 1; if not, goes to next function body and calls itself again.

3.2.2 Guards

Patterns are used to specify data should form to some form and deconstruct it, whereas guards are used to check whether some properties of an argument are true or false. It is similar to an if statement, and more likely to a ‘switch’ block in Java. Following is an example of using guards:

```
max :: (Ord a) => a -> a -> a
max a b | a > b      = a
        | otherwise = b
```

The function above returns the maximum value between the inputs *a* and *b*. Notice that, ‘Ord’ here enables ordering on a general type *a*, so we can apply ‘>’ ‘<’ to compare two variables in type *a*.

3.2.3 Local Definition

The ‘where’ block is used to simplify complex codes with newly defined variables. It avoids repeated expressions: an expression is only calculated once and it is binded to a name than then the name can be referred to instead of the whole expression.

```
divA :: Float -> Char
divA a | a/2 <= 10 = 'A'
        | a/2 <= 20 = 'B'
        | a/2 <= 30 = 'C'
        | otherwise = 'D'
```

The function above firstly calculate the input parameter *a* divided by 2, and decides which sector it is in and returns a character. It repeated defines the expression *a/2* for three times. We can avoid that by using ‘where’ statement as below:

```

divA :: Float -> Char
divA a | b > 10  = 'A'
      | b > 20  = 'B'
      | b > 30  = 'C'
      | otherwise = 'D'
      where b = a/2

```

As we can see, it improves the readability by giving names to an expression and can accelerate programs since expressions like $a/2$ is calculated only once. Mention that, these names like b are visible across the guards.

3.2.4 Self-defined Types

We can also create our own data types in a very straight forward way. One way is to use the **data** keyword.

```
data Bool = False | True
```

data means that we are defining a new data type. The part on the left of $=$ denotes the type name, which is **Bool**. The parts after the $=$ are value constructors. They specify the different values that this type can have. The symbol $|$ stands for ‘or’, so the type **Bool** can either be **False** or **True**.

Following is another example that defines a data type **Shape**

```
data Shape = Rectangle Float Float | Circle Float | Triangle Float Float Float
           deriving (Show, Eq)
```

A **Shape** can either be a **Rectangle** or a **Circle** or a **Triangle**. The **Rectangle** value constructor has two fields which are floats. Fields after a value constructor are basically the parameters that type contains. Here, the two fields of **Rectangle** are length and width. Similarly, **Circle** has one field which is the radius. **Triangle** has three fields which stands for the length of its three edges. The ‘deriving’ statement after the data type declaration enables the defined data type to be shown and to be compared to check if two shapes are the same. For example, ‘Circle 2.4 == Circle 2.5’ would return false.

3.3 Model the λ -calculus and the Curry Type Assignment System in Haskell

Since the lambda calculus lies at the basis of functional languages like Haskell. Implement lambda calculus in Haskell is quite a straight forward work. In this project, both the typed and untyped lambda calculus are implemented. In addition, the λ xgc-calculi is also implemented based on the ordinary λ -calculus. The implementation consists three main parts: data types declaration, parser, and the main mechanism.

3.3.1 Data Types

The ordinary λ -calculus data types have already been introduced in Section 1.5. Based on that, the λ xgc-terms are defined as following:

```

type Name = Char
data Term = Var Name | Abs Name Term | App Term Term | XSub Term Term Term
           deriving (Show, Eq)

```

There is an additional term called **XSub** which stands for the explicit substitution terms of the form $M\langle x := N \rangle$. It is worth mentioning that if a λxgc -term does not contain any subterms of the form $M\langle x := N \rangle$, then it is ‘pure’. The abstraction in λxgc -calculus is different from the ordinary λ -calculus: an abstraction in λxgc -calculus can also in the form $\lambda x.M\langle x := N \rangle$.

The data types of the Curry type assignment system are defined as following:

```
data Type = TP Type Type | TVar [Char]
           deriving (Show, Eq)

type Predicate = Type           --redundant
type Subject = Term            --redundant
type Context = [Statement]
type Statement = (Subject, Predicate)
type Derivation = (Context, [Statement], Statement)
type PPc = (Context, Type)

data TSub = Sub (Type, Type) | So (TSub, TSub)
           deriving (Show, Eq)
```

A type can be either a single type such as A, B, \dots or formed by several types $(A \rightarrow B) \rightarrow C$. So the data type **Type** can either be a **TP** or **TVar**. The ‘deriving’ statement enables this data type can be printed and compared with each other. Furthermore, there are several other concepts in Curry type assignment system such as **predicate**, **subject**, **context**, **statement**. The **predicate** and **subject** definition is redundant since we can directly use **Type** and **Term** to represent them. The type **PPc** is created to represent the result generated by principle pair algorithm, which contains a context and a type assigned to the input term. A type substitution

3.3.2 Parser

To allow interactions and inputs from the users, the program needs syntax rules to define how information should be written and it needs a parser to parse the input string. Following is the parser in Haskell implementation. It only allows simplified term that parentheses and repeated λ are omitted. So the term $((\lambda x.\lambda y.y)(\lambda z.z))(\lambda w.w)$ should be input as $(\lambda xy.y)(\lambda z.z)(\lambda w.w)$.

```
to_term :: [Char] -> Term           --parse a string into lambda term
to_term (a:exp) | a == '/' = to_abs exp -- / as lambda.
                                -- Abstraction starts with /
to_term [a] = to_var [a]           --single character, variable
to_term exp = to_app exp           --otherwise, application

to_var :: [Char] -> Term           --parse a character into variable
to_var [a] = Var a

to_abs :: [Char] -> Term           --parse a string into abstraction
to_abs ('.' :rest) = if tail rest == [] then to_var rest
                    --if rest is a single character, then parse it into variable
                    else to_app rest
                    --otherwise, it is an application
to_abs (a:rest) = Abs a (to_abs rest)

to_app :: [Char] -> Term           --parse the function and argument
to_app exp = App (to_term lhs)(to_term rhs)
              where (lhs, rhs) = div_sub exp

div_sub :: [Char] -> ([Char], [Char]) --divide a term into 2 parts
```

```

div_sub expr | last expr == ')' && last lhs == ')' && head (tail expr) == '/'
    = (to_expr lhs, rhs)
| last expr == ')'
    = (lhs, rhs)
| head (tail expr) == '/' && last (init expr) == ')'
    = (to_expr (init expr), [last expr])
| otherwise = (init expr, [last expr])
    where (lhs, rhs) = split (init expr) [] 1

to_expr :: [Char] -> [Char]    --extract the expression from a string
to_expr ex | last ex == ')' && lhs == [] = rhs
    | otherwise = ex
    where (lhs, rhs) = split (init ex) [] 1

split :: [Char] -> [Char] -> Integer -> ([Char], [Char])
{-|
    it marks the rightmost ')' and finds the matching '(' and extract the term out
    without brackets
-}
split left right n
    | n == 1 && last left == '(' = (new_left, right)
    | last left == '(' = split new_left new_right (n-1)
    | last left == ')' = split new_left new_right (n+1)
    | otherwise = split new_left new_right n
    where new_left = init left
          new_right = (last left : right)

```

3.3.3 Main mechanism

The main mechanism of λ -calculus, λ xgc-calculus, and Curry type assignment only contains 300 lines of code in Haskell. The full implementation code can be found in Appendix A.

3.4 Issues and Weakness of Implementation in Haskell

The functional programming languages are elegant and concise. The implementation of untyped lambda calculus, explicit substitution, and Curry type assignment system only contains several hundred lines of code in Haskell. Each function is defined by pattern matching and readable. It clearly shows what each function do according to the input pattern. Local definitions are utilized to avoid duplicate execution on expressions, it makes the program run faster. However, it has some issues; and as an application or software, it has less usability.

Chapter 4

The λ -calculus and Curry Type Assignment System Implementation in Java

The Java version of implementation is transferred directly from Haskell. However, the semantics and operational order of Haskell is difference from Java, especially the ‘where’ statement of Haskell. The ‘where’ declaration is called when the defined variables are used, however, in Java, we need to separate a ‘where’ block into parts according to where it is called. An example is shown in Section 4.1. This difference is worth taking care of when transfer the Haskell into Java, otherwise, `NullPointerException` may occur.

In all data types, the method `toString()` has been implemented – more precisely overridden, as it is a method of the Java primary class `Object`. This function is used to print out the data type as a string. There are more method signatures defined in the interfaces as a reference to instantiable classes.

4.1 Different operational order between Haskell and Java

As mentioned above, the operational order of Haskell is different from Java, not only the pattern matching, but also the ‘where’ declarations. Since the code in Haskell cannot be transferred into Java directly, it largely abandoned the simplicity of codes. Following is a piece of code that defines principle pair algorithm:

```
ppc :: Term -> [Char] -> (PPc, [Char])
ppc (Abs x y) r | contains (Var x) pi = ((removeItem (Var x, a) pi, (TP a p)), t11)
      | otherwise = ((pi, (TP f p)), t11)
                        where (f, t1) = fresh r           (1)
                        ((pi, p), t11) = ppc y t1         (2)
                        (_, a) = search (Var x) pi         (3)
```

‘Guards’ are used as an if-then-else block. There are three declarations in the ‘where’ block each with a distinct number. For the first condition of `ppc` function, declaration (1)(2)(3) are called since all the declared variables are used. In the ‘otherwise’ condition, only declaration (1)(2) are called. In this case, in order to transfer the Haskell function into Java methods, we need to repetitively define the variables which largely reduces the simplicity of codes.

```
public PPC ppc(Term term, String counter){
    ...
    else if(term instanceof Abstraction){
        Variable xv = new Variable(((Abstraction) term).getName());
        PPC receiver = ppc(((Abstraction) term).getTerm(), counter);
```

```

    if(receiver != null){
        if(contains(xv, receiver.getSubject())){
            Type searchType = search(xv, receiver.getSubject()).getPredicate();
            ArrayList<Statement> original = receiver.getSubject().getContext();
            ArrayList<Statement> remove = new ArrayList<>();

            for(Statement s: original){
                if(s.getSubject().equals(xv)) remove.add(s);
            }

            for(Statement rm: remove){
                original.remove(rm);
            }

            TP tp = new TP(searchType, receiver.getPredicate());
            return new PPC(new Context(original), tp, receiver.getCounter());
        }
        else{
            TVar f = new TVar(receiver.getCounter().substring(0, 1));
            return new PPC(receiver.getSubject(), new TP(f, receiver.getPredicate()),
                           receiver.getCounter().substring(1));
        }
    }
    else {
        return null;
    }
}
...
}

```

4.2 λ -calculus Representation

4.2.1 λ -terms

A Lambda term can either be a variable, an abstraction or an application. An interface ‘Term’ is defined as a reference type, which has three lambda term implementations. This is the generic representation of all lambda terms.

This interface defines the method `equals(Term t)`, which return a boolean states whether it is equal to the input term. This method overrides the `equals(Object o)` method of the Java primary class `Object`. It is essential in the reduction procedure, since we can decide whether a term is reducible by compare the term before reduction and after. If the term is the same as it before reduction, then it cannot be reduced. Therefore, we know when we can stop the reduction procedure.

As mentioned in Section 1.5, there are five reduction strategies enabled in the reducer. So there are five method signatures defined in the interface: `evaluateNormal(Bool exsub)`, `evaluateCbn(Bool exsub)`, `evaluateCbv(Bool exsub)`, `headReduction(Bool exsub)`, `applicativeOrder(Bool exsub)`. Each of these method refers to a specific reduction strategy as the method name.

The method `mirror()` is used to create a mirror term with exactly the same class variables. This method is used when the substitution is performed. For example, if we have a lambda term $\lambda f.(\lambda x.f(xx))(\lambda x.f(xx))$, it is reduced to $\lambda f.f((\lambda x.f(xx))(\lambda x.f(xx)))$. If we do not create a mirror term of $(\lambda x.f(xx))$, both of these two $(\lambda x.f(xx))$ would refer to the same memory space. In other words, the abstraction $(\lambda x.f(xx))$ is used twice to form an application. Therefore, if operations performed on the function, it is also performed on the argument. It is essential to create a mirror

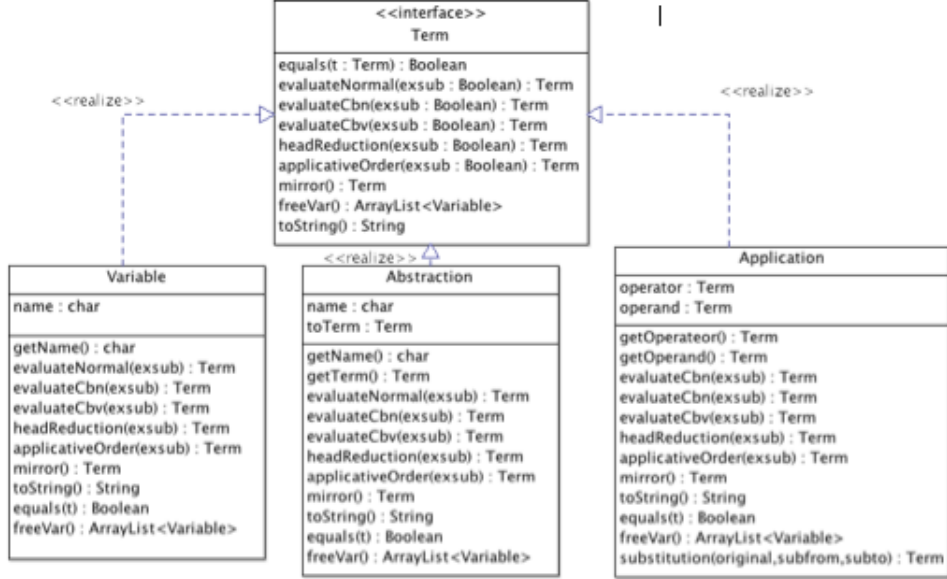


Figure 4.1: Class diagram of Lambda Calculus data structures

term that separate those two terms although they look exactly the same.

The method `freeVar()` is used to get all the free variables of a λ -term. It is used to perform α -conversion when we substitute bound variables in an abstraction. The free variables are defined in Definition 1.1.

Finally, the `toString()` method of class `Object` is overridden to transform a lambda term to a string.

Notice that, in the implementation class `Application`, there is a `substitution(Term original, Term subfrom, Term subto):Term` method. It is used to substitute bound variables in an abstraction, which is an application function, to the argument. The first parameter of the method is the body of abstraction, the second argument states which bound variable could be substituted and the third argument states what it could be substituted to.

4.2.2 Interpreter

To allow interactions and inputs from users, the program needs to interpret the input term into a data structure in the system. Figure 4.2 illustrates all the methods that used to interpret an input into a lambda term data structure.

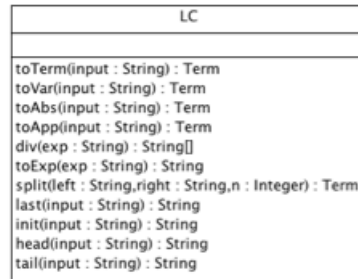


Figure 4.2: Interpretation methods in LC class

The method `toTerm(String input)` takes a string as input and returns a `Term`. It calls `toVar(String input)`, `toAbs(String input)` or `toApp(String input)` according to type of the outermost term. Then, those three methods iteratively call each other until the whole term is interpreted.

`div(String exp)` divides an application into the function and argument. Basically, it return an array with 2 string elements. The method `toExp(String exp)` is used to remove brackets of

applications and returns an expression that is bracker-free. Finally, the most important method `split(String left, String right, Int n)` implements the bracket removal mechanism. It marks the right most right bracker ')' and find the matching left bracket '(' and extract the term out without brackets.

Since the Java implementation is directly transferred from Haskell, there are also four basic string operations which are Haskell built-in functions. `last(String input)` takes a string and returns the last character of the string, it returns itself if the length of input string is 1. `init(String input)` accepts a string and returns the string without its last character, it returns an empty string if it only contains one character. `head(String input)` takes a string and returns the first element of the string, it returns itself when the input string length is 1. `tail(String input)` takes a string and returns the string without its first element, it returns an empty string when the length of input is 1.

Following is the example of how those Haskell built-in function work:

<code>init "Hello"</code>	<code>"Hell"</code>	<code>head "Haskell"</code>	<code>"H"</code>
<code>init "A"</code>	<code>""</code>	<code>head "A"</code>	<code>"A"</code>
<code>last "String"</code>	<code>"g"</code>	<code>tail "Lambda"</code>	<code>"ambda"</code>
<code>last "A"</code>	<code>"A"</code>	<code>tail "A"</code>	<code>""</code>

4.3 Curry Type Assignment System Representation

4.3.1 Types

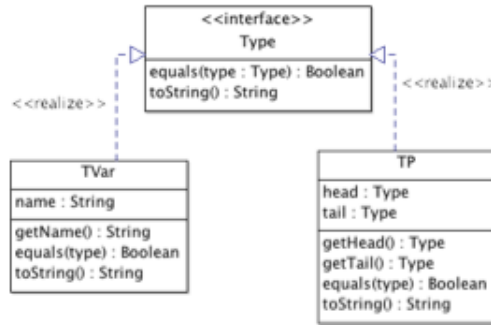
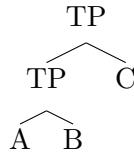


Figure 4.3: Class diagram of type representation

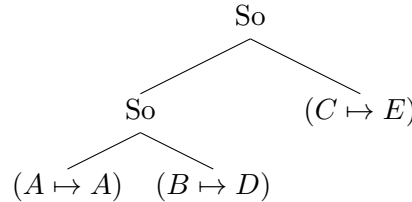
Type assignment system assigns types to λ -terms. In \mathbb{T} , a type can either be a single type(A, B, \dots) or a complex type($A \rightarrow B$). Since the complex type is formed by single types, it can be represented using the binary tree structure, where single types are leafs and complex types are branches. For example, the type $(A \rightarrow B) \rightarrow C$ is represented as `TP(TP (TVar A)(TVar B))(TVar C)`, or in a more intuitionistic way:



Illustrated in Figure 4.3, the interface `Type` is defined as a reference type. It defines two basic method signatures: `equals(Type type)` and `toString()` which override the primary methods in `Object`. There are two implementation class: `TVar` and `TP` which stand for single type and complex type. `TVar` contains a class variable `name` which states the type name. It implements method signatures defined in the interface and a `get` method that returns the type name. The implementation class `TP` is similar to a `Node` of a tree structure, which contains two class variables: left branch and right branch. Method signatures and `get` methods are implemented in the class.

4.3.2 Type substitution

Principal types for λ -terms are defined using the notion of unification of types that defined by Definition 2.5. Similar to types structure in Section 4.3.1, a type substitution can either be a single substitution ($\varphi \mapsto B$) or a complex substitution ($S_1 \circ S_2$). It could also be represented as a binary tree structure. For example, when we unify $(A \rightarrow B) \rightarrow C$ and $(A \rightarrow D) \rightarrow E$, we would get the complex substitution $((A \mapsto A) \circ (B \mapsto D)) \circ (C \mapsto E)$ which also can be represented as a binary tree:



As shown in Figure 4.4, an interface `TSub` is needed as a reference type. It doesn't have any method signatures, operation on substitutions are defined as auxiliary function in class `LC`.

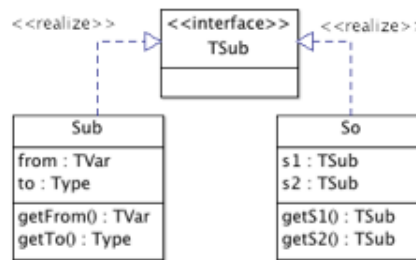


Figure 4.4: Class diagram of type substitutions representation

4.3.3 Other data types

The Curry type assignment system defines other concepts besides types and substitutions. Statement is essential to describe the type assigned to a λ -term; Context is used to collect all statements used for free variables of a term when typing that term; and PPC is a data type that stores context and the type assigned to that term typed.

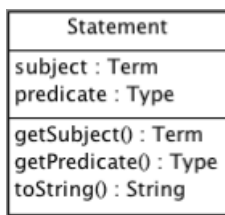


Figure 4.5: Class diagram of Statement

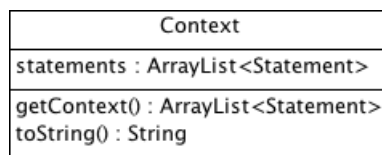


Figure 4.6: Class diagram of Context

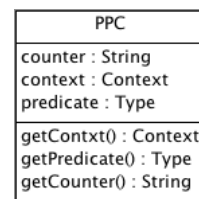


Figure 4.7: Class diagram of PPC

The `Statement` class has two attributes which are subject and predicate. It contains basic get method that return private attributes, and it also overrides the `toString()` method to transform a `Statement` into a string as "subject : predicate"

The `Context` class only contains an `ArrayList` of `Statement` as attribute. A context could have none statement if all the variables are bound. The `toString()` method returns the context a string in the format: "statement1, statement2, ..."

The `PPC` class, storing principle pair algorithm results, is a little more complex. As defined in Definition 2.6, the `ppc` function return a context and the type assigned to the term typed as $\langle \text{context}, \text{type} \rangle$. A `PPC` instance, stores the result of `ppc` function, therefore it contains two attributes as context and type. In additional, it also has a counter attribute. Since principle pair

algorithm always create fresh types, we should keep track of what type names are available that have not been used. It basically states all the available type names.

4.3.4 Core type assignment mechanism implementation

The Curry type assignment implementation is built based on the lambda calculus implementation. As it extends more operation on λ -terms for type assignment, there are more auxiliary methods defined in class LC.

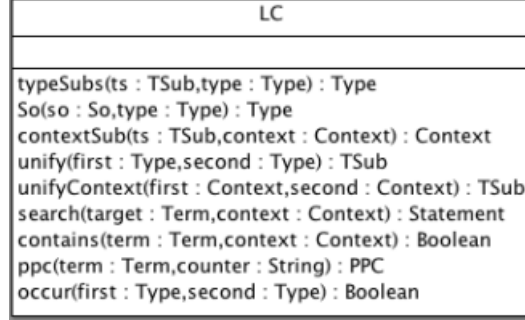


Figure 4.8: Class diagram of Lambda Calculus data structures

The methods listed in Figure 4.8 are the core type assignment mechanism implementations. The top-level method is `ppc(Term term, String counter)`, when a term comes in and needs to be typed, the `ppc` method is called and it iteratively calls other auxiliary methods. Finally, it will return a `PPC` instance which contains a context and the typed assigned.

Given a type substitution and a type, the method `typeSubs(TSub, Type)` substitutes all the type variables according to the substitution. It implements the type substitution defined in Definition 2.4. Notice that, since a `TSub` can either be a single substitution `Sub` or a complex `So` as defined in Section 4.4, it calls `So(So, Type)` method that implements the `So` substitution described in Definition 2.4 (b).

The method `contextSub(TSub, Context)` implements the context substitution defined in Definition 2.5 (c).

The method `unify(Type, Type)` implements the type unification algorithm defined in Definition 2.4. And `unifyContext(Context, Context)` implements the context unification in Definition 2.5 (ii). It should be mentioned that, since Id_s stands for the substitution that replaces all type variables by themselves, it could be simplified as a substitution that only contains the substitution $(A \mapsto A)$. We do not care what the type A is and whether it is a valid type variable, because the substitution $(A \mapsto A)$ would not affect any unifications.

The `search(Term, Context)` method is an auxiliary method that given a λ -term and a context, it returns the corresponding statement of that term in the context. This method is used in two places. Firstly, it is used when we want to unify the context Γ_0 and Γ_1 . We need to find the statements with the same subject in these two context and unify them. Therefore, when we traverse through the context Γ_0 , the search method is necessary to find the corresponding statement in Γ_1 . Secondly, it is used in the principal pair algorithm when we type an abstraction. The temporary type assigned to bound variables should be removed from the context. Therefore, given a variable, the search method can find the corresponding statement and further be removed.

The method `contains(Term, Context)` takes a λ -term and a context as arguments and return a boolean states whether the context contains a statement for the term.

`ppc(Term, String)` is the core principal pair algorithm implementation method. The definition of principal pair algorithm can be found in Definition 2.6. The second argument in `String` type is the counter of all available type names.

Finally, the method `occur(Type, Type)` is used to decide whether a type occurs in another. It is used in the second type unification algorithm defined in Definition 2.5.

4.4 Explicit substitution and garbage collection

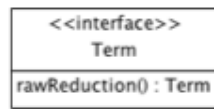


Figure 4.9: Class diagram of Lambda Calculus data structures

4.5 Syntax

Chapter 5

GUI development

Haskell has at least four toolkits for programming a graphical interface: wxHaskell, Gtk2Hs, hoc, and qtHaskell. However, Java Graphics APIs—AWT and Swing—provide a huge set of reusable GUI components, such as button, text field, label, panel and frame for building GUI applications. So, it is more flexible to develop a GUI by reusing these classes in Java. I believe it would also be an interesting work to implement a comprehensive GUI in Haskell. In this project, a useable and functional GUI is developed in Java. The GUI design is based on Sestoft’s ‘lambda term reducer’ which could be found at “<http://www.itu.dk/people/sestoft/lamreduce/>”, and the draft report describing the implementation in Sestoft[11]. Some improvements have been made such as to deal with non-terminating terms in trace enviroment. In addition, the Curry type assignment system is also enable in the GUI which is run in parallel with λ -term reduction.

5.1 Layout

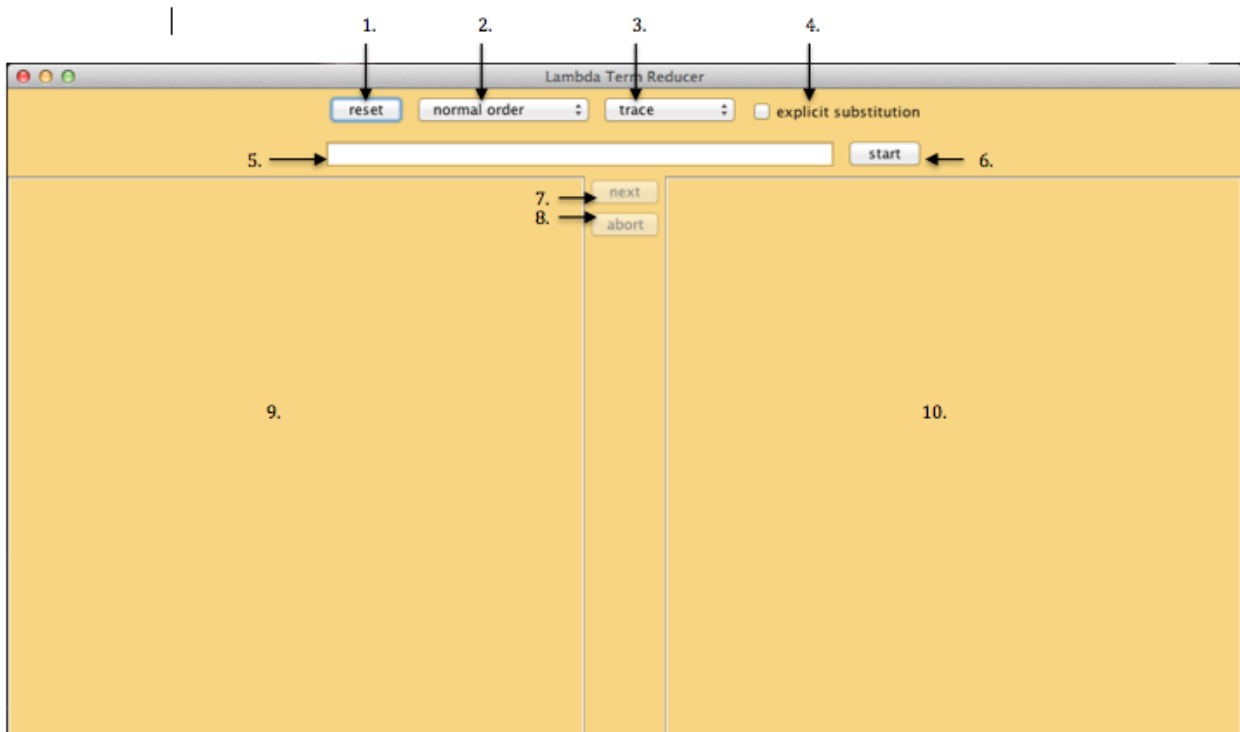


Figure 5.1: Layout of the Graphical User Interface

The function of each button or text area label in Figure 5.1:

1. reset button: reset all the button, drop-down lists values to default and empty the output areas 9 and 10.

2. reduction strategy drop-down list: 5 reduction strategies can be selected by the user.
3. reduction step drop-down list: the user can select how the reduction is performed: normalize: directly output the reduction result by a specific reduction strategy and states how many β -reductions performed. trace: output the reduction procedure for each β -reduction performed. single step: it only outputs the reduction of a single step, moreover, the buttons 7 and 8 are enabled to let user choose whether continue reduction or abort the procedure.
4. explicit substitution check box: user can tick to enable explicit substitution and garbage collection.
5. input field: user input the term to be reduced and typed in this field.
6. start button: press the button to start the procedure with pre-specified reduction strategy(2) and reduction step option(3).
7. next button: carry on reduction when single step is selected.
8. abort button: abort the reduction when the input term is non-terminate or in the middle of single step reduction.
9. reduction output area: the output of reduction procedure is printed out in this area. The format is decided by the reduction step option(2).
10. type assignment output area: the output of type assignment result is printed out in this area. For normalize and trace, only the input term is typed. However, if the single step strategy is selected, intermediate terms during reduction would also be typed.

5.2 Functionalities

In this section, the functionalities of the GUI would be shown by some sample runs. We mainly test the difference in different tracers: normalize, trace, and single-step. Moreover, the improvements compared to Sestoft's lambda term reducer would be shown by example. Here, we will not give any examples on the differences between outputs by different reduction strategies, since it has already been shown in Chapter 1.

5.2.1 'Normalize' tracer

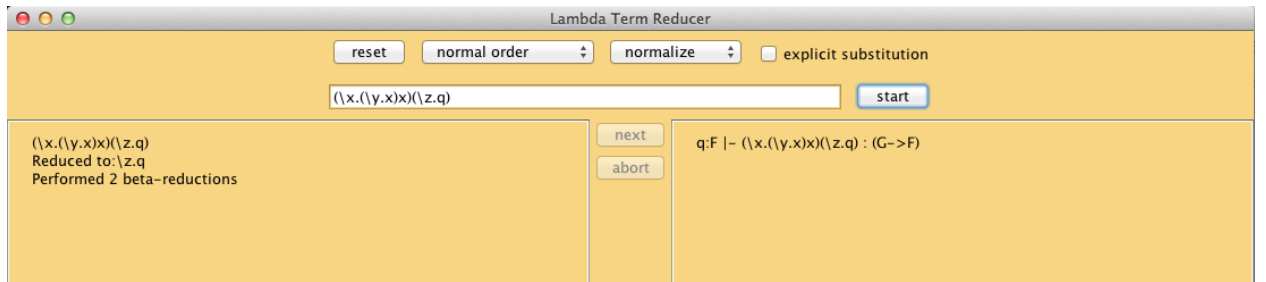


Figure 5.2: Sample run in normalize tracer

When 'normalize' tracer is selected, the reducer tries to normalize the input λ -term with a specific reduction strategy. In other words, it reduces the input term with a reduction strategy until no more β -reduction can be performed. The intermediate terms and the reduction procedure would not be shown in this case. It outputs the final term reached and displays the total β -reduction performed. On the right output area, the type assignment result is displayed. It is displayed in the derivation format: context \vdash statement, which means the statement is derivable from the context. Recall that, the context is used to collect all statements used for the free variables of a term when typing that term. In this example, only q is free others are all bound.

5.2.2 ‘Trace’ tracer

The tracer ‘trace’, as its name, trace over the reduction procedure and display every reduction steps, intermediate terms and corresponding type assigned to that term.

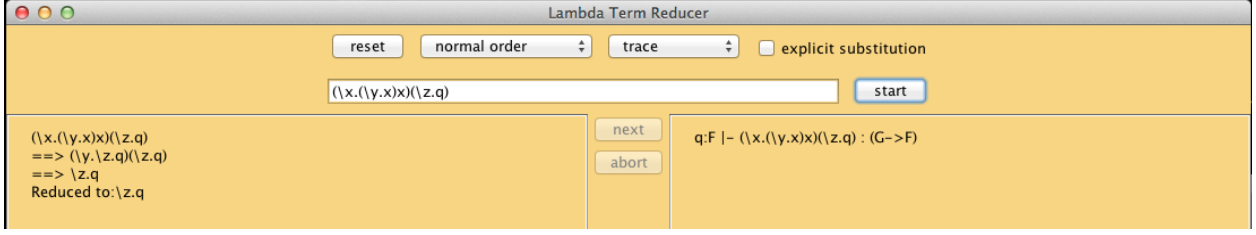


Figure 5.3: Sample run using ‘trace’ tracer

5.2.3 ‘Single-step’ tracer

The ‘single-step’ tracer performs β -reduction once per time and user can control whether continue or abort.

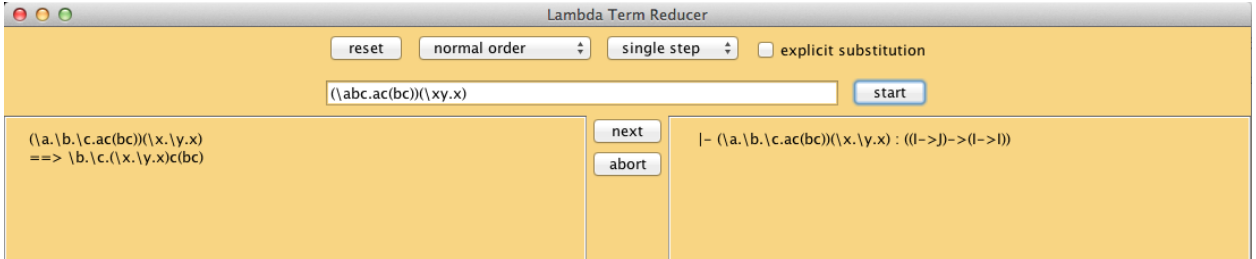


Figure 5.4: Sample run in single-step tracer

As shown in Figure 5.4, after the ‘start’ button is pressed, a single β -reduction is performed and the corresponding type assignment result is displayed. As we can see, the ‘next’ button and ‘abort’ button are enabled now. The user can press the ‘next’ button to perform another single β -reduction step, or press the ‘abort’ button to abort the whole procedure.

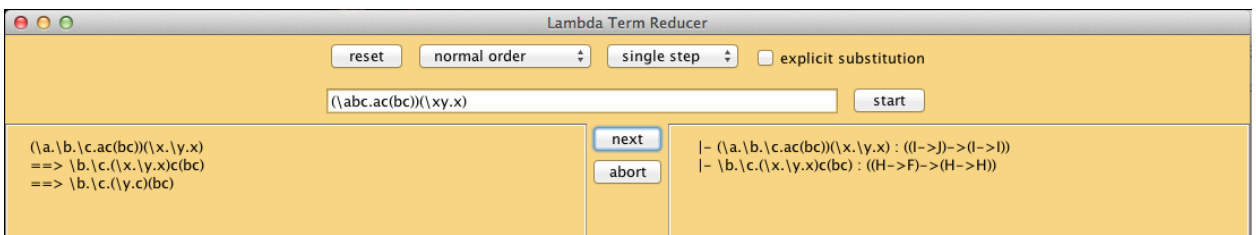


Figure 5.5: Sample run by single-step tracer–second step

Each time the ‘next’ button is pressed, a single β -reduction would be performed until no reduction can be performed according to the reduction strategy. In Figure 5.5, another β -reduction is performed and the corresponding type assignment result is displayed.

When there is no β -reduction to perform on the λ -term, the reduction procedure stops and both ‘next’ and ‘abort’ buttons are disabled.

The ‘single-step’ tracer in Sestoft’s implementation is slightly different from this implementation. As shown in Figure 5.7, Sestoft’s reducer highlights the redex in the term and let the user to click on the redex to perform next reduction step. The highlight of redex in the term is useful since it shows which redex the reducer chooses to reduce next. However, it gives less details of the reduction as a complete procedure. The user may not remember the reduction procedure since previous intermediate terms are not displayed. Therefore, in this project, the whole procedure is shown and

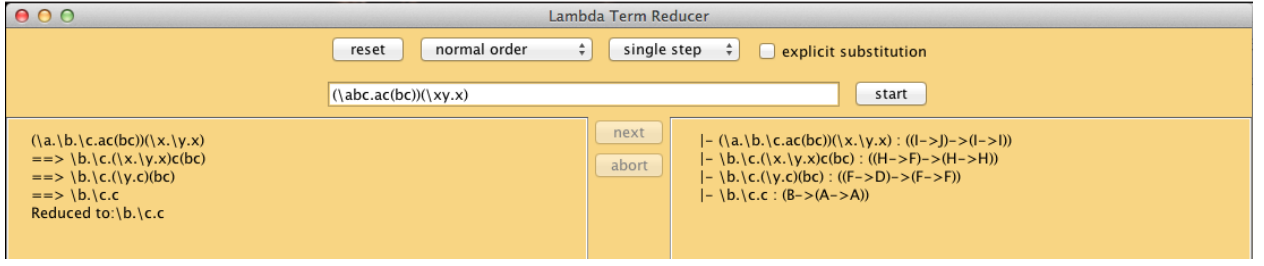


Figure 5.6: Sample run in single-step tracer—final result

each single-step reduction adds a new line in the procedure. The final output of reduction is the same as ‘trace’, except that each intermediate terms are also typed.

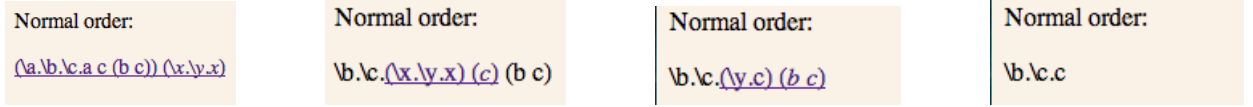


Figure 5.7: Sestoft’s lambda term reducer using ‘single-step’ tracer

5.2.4 Deal with non-terminating terms by trace

In Sestoft’s implementation, the engine enters an infinite loop and stuck when the input term is non-terminating and the tracer is ‘trace’. There is no intermediate terms or steps displayed, so the user would not know what is happening and what the problem is. In this project, when the input term is non-terminating and the tracer is ‘trace’, all the intermediate steps would be displayed. As we can see from Figure 5.8, the left output area is a scroll pane, where the user can scroll down to view the whole procedure. Notice that, the input term $\backslash f.(\backslash x.f(xx))(\backslash x.f(xx))$ is untypable since there is a self-application xx , the reason why it cannot be typed could be found in Example 2.1.1.

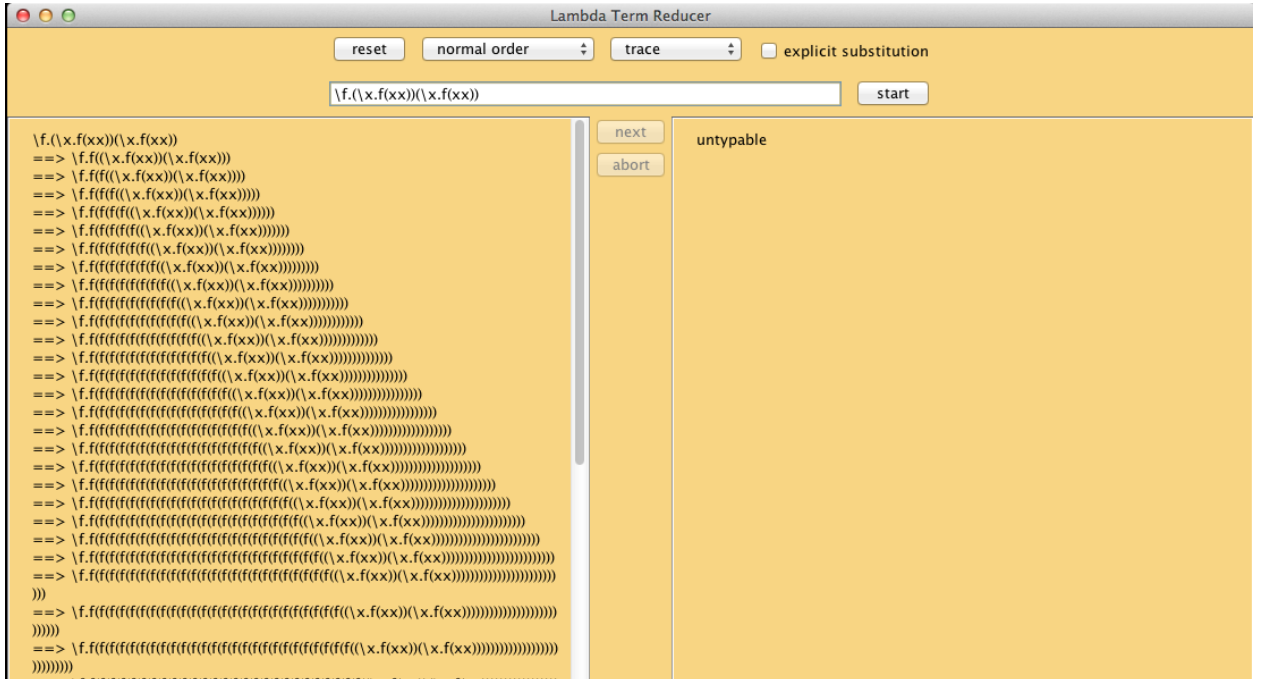


Figure 5.8: Sample run of a non-terminate term by trace

5.2.5 Explicit Substitution

Figure 5.9 illustrates the output when explicit substitution is enabled.

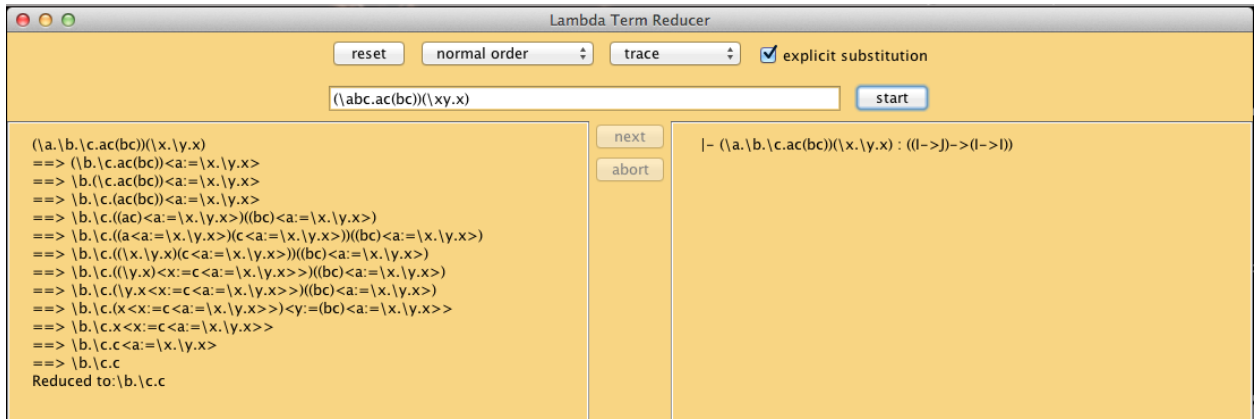


Figure 5.9: Sample run when explicit substitution enabled

Chapter 6

Java vs Haskell, Pros and Cons

6.1 Pros

6.2 Cons

Chapter 7

Conclusion

Bibliography

- [1] Hendrik Pieter Barendregt. *The lambda calculus*, volume 3. North-Holland Amsterdam, 1984.
- [2] Roel Bloo and Kristoffer H Rose. Preservation of strong normalisation in named lambda calculi with explicit substitution and garbage collection. In *In CSN-95: Computer Science in the Netherlands*. Citeseer, 1995.
- [3] Alonzo Church. A formulation of the simple theory of types. *The journal of symbolic logic*, 5(02):56–68, 1940.
- [4] Haskell B Curry. Functionality in combinatory logic. *Proceedings of the National Academy of Sciences of the United States of America*, 20(11):584, 1934.
- [5] Haskell B Curry, J Roger Hindley, and Jonathan P Seldin. *Combinatory Logic: Vol.: 2*. North-Holland Publishing Company, 1972.
- [6] Haskell Brooks Curry, Robert Feys, William Craig, J Roger Hindley, and Jonathan P Seldin. *Combinatory logic*, volume 2. North-Holland Amsterdam, 1972.
- [7] Chris Hankin. *Lambda calculi: a guide for computer scientists*. 1994.
- [8] Miran Lipovaca. *Learn you a Haskell for great good!* No Starch Press, 2011.
- [9] Lawrence C Paulson. *ML for the Working Programmer*. Cambridge University Press, 1996.
- [10] Raul Rojas. A tutorial introduction to the lambda calculus. FU Berlin, WS-97/98, 1997.
- [11] Peter Sestoft. Standard ml on the web server: Visualizing lambda calculus reduction. 1996.
- [12] Peter Sestoft. Demonstrating lambda calculus reduction. In *The essence of computation*, pages 420–435. Springer, 2002.
- [13] Steffen van Bakel. Type systems for programming language, course notes. 2001.