



University of Glasgow | School of  
Computing Science

# Building a Lambda Expression Evaluator as a Learning Support Tool

Yola Jones

2409217J

Supervisor: Ornela Dardha

School of Computing Science  
Sir Alwyn Williams Building  
University of Glasgow  
G12 8QQ

A dissertation presented in part fulfilment of the requirements of the  
degree of MSc in Software Development at the University of Glasgow

September 2019

## **Abstract**

This dissertation describes the design, implementation and evaluation of a lambda calculus expression evaluator built using ANTLR and contained within a web interface, designed as support material for the Theory of Computation course at the University of Glasgow. This dissertation discusses in detail the relevant sections of lambda calculus as covered on the course, then discusses the process of using Abstract Syntax Trees to examine the structure of a term based on a particular syntax, which can then be used to parse and evaluate input expressions. The approach taken to evaluation using ANTLR has also been discussed in detail, before giving a summary of testing done and evidence that the goals of the project have been met.

There were two key goals for this project, to design an interface which would accurately evaluate lambda terms, and which would help improve students understanding of those on the Theory of Computation course. Both of these goals were met, with testing done to prove the accuracy of lambda term evaluation and a user test carried out in order to determine the effectiveness of the interface in improving understanding. Positive results were gained both from qualitative user feedback and statistical analysis using quantitative data taken from these tests.

The code for this project has been submitted alongside this dissertation, but can also be found at <https://github.com/jonesy30/LambdaCalculator> on Github. This tool will be used in future years to support those on the Theory of Computation course.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>2</b>
2.1	Lambda Calculus . . . . .	2
2.1.1	Basic Principles . . . . .	2
2.1.2	Alpha Conversion . . . . .	3
2.1.3	Reduction . . . . .	3
2.1.4	Typing Rules . . . . .	4
<b>3</b>	<b>Analysis</b>	<b>5</b>
3.1	Lambda Calculus . . . . .	6
3.2	Abstract Syntax . . . . .	6
3.2.1	ANTLR . . . . .	6
3.3	Interface . . . . .	8
<b>4</b>	<b>Aims and Objectives</b>	<b>8</b>
<b>5</b>	<b>Design and Implementation</b>	<b>9</b>
5.1	Design . . . . .	9
5.1.1	Grammar . . . . .	9
5.1.2	Expression Evaluator . . . . .	9
5.1.3	Web Interface . . . . .	10
5.2	Implementation . . . . .	11
5.2.1	Grammar . . . . .	11
5.2.2	Abstract Syntax Tree . . . . .	11
5.2.3	Expression Evaluator . . . . .	11
5.2.4	Web Interface . . . . .	16
5.2.5	Complete Program . . . . .	16
<b>6</b>	<b>Testing</b>	<b>16</b>
6.1	Accuracy of Lambda Term Evaluation . . . . .	16
6.2	Interface Usefulness and Usability . . . . .	20
6.2.1	Testing . . . . .	20
6.2.2	Results . . . . .	20
<b>7</b>	<b>Evaluation</b>	<b>23</b>
7.1	Accurately Evaluating Lambda Terms . . . . .	24
7.1.1	Limitations . . . . .	24
7.2	Improving Student Understanding . . . . .	24
7.3	Future Work . . . . .	24
<b>8</b>	<b>Conclusion</b>	<b>25</b>

# Chapter 1: Introduction

Lambda calculus is often reported as being the worlds smallest programming language [22] expressing terms as mathematical functions which can represent any computable problem [27]. Students on the Theory of Computation course taught at The University of Glasgow learn lambda calculus to give them an understanding of functional programming and other programming paradigms [11].

This report will cover the work done in producing a web application which will verify and evaluate lambda calculus terms input by a user, in order to support future students learning functional programming. The project has been split into three major tasks:

1. Understand and use ANTLR to create a parser which will read lambda terms and produce an abstract syntax tree from the user input
2. Write a program which will use this tree to find the result of the incoming term
3. Create a web interface which will allow users to input lambda calculus expressions and have them checked and evaluated, in order to help improve their understanding of lambda calculus

Since the focus is the Theory of Computation course, learning outcomes and course material have been analysed in order to produce a tool which accurately reflects what is taught. An effort has been made however to make the tool accessible to those learning lambda calculus through different means, for example self teaching or those at other universities. This has resulted in an interface which is aimed at supporting the lecture material but not limited by it, with goals of future work aimed at expanding the interface further to support alternative notation than what is covered in the lectures.

Overall the project was a success, an interface was created which accurately evaluates lambda terms through a web interface. The tool implements simply typed applied lambda calculus and allows for different reduction strategies for those who want a more in depth understanding, wrapped in a simple and clear layout that is easy to use. User testing found that participants enjoyed the interface and said they would use such an interface if they were learning lambda calculus, reporting a significant increase in understanding having used the interface for a short time which is verified by statistical analysis. The tool created for this project will be given to students on the Theory of Computation course when it is next taught at the University of Glasgow, in the hopes that it will aid understanding by giving students practical support.

The following report documents the process in creating such a tool, by introducing key concepts and literature and refining specific project objectives, then expanding on the design and implementation stages, before finally documenting the testing process and evaluating whether or not the goals of the project have been met.

# Chapter 2: Background

## 2.1 Lambda Calculus

Lambda calculus is a formal syntax for expressing computation statements as mathematical expressions [27] created by Alonzo Church in 1936 [2]. It is an extremely simple model of computation, while still being Turing complete [26] and forms the basis for functional programming languages like Haskell [2]. Pure lambda calculus terms are comprised of three elements, variables (such as  $x$  or  $y$ ), abstractions (with one input and one output, in the form  $\lambda x.M$ ) and applications (in the form  $MN$ ) [12].

Applied lambda calculus expands this notation, adding constants which act as values and operations [28]. Typed lambda calculus extends this further, adding types to each term, allowing us to restrict the operations available to different types of objects and ensuring that our expression is runnable [20]. The combination of applied and typed lambda calculus gives us a grammar which can be used for type checking and arithmetically evaluating lambda calculus terms.

The Theory of Computation course teaches simply typed lambda calculus (applied lambda calculus with types), and includes lectures on alpha conversion and beta reduction, alongside the definition of normal form and typing rules. As a result, these will be covered in the following sections in an attempt to use lecture material to define requirements of the final system.

### 2.1.1 Basic Principles

As discussed above, the lambda grammar is made up of three elements: variables, abstractions, and applications. Abstractions are representations of single input functions which take one input, and use this input to compute one output. An example abstraction  $(\lambda x.x + 1)$  is equivalent to the function  $f(x) = x + 1$ , taking one input  $x$ , and returning the value of  $x$  incremented by 1. Multi-input functions can be expressed in nested abstraction terms, for example  $f(x, y) = x + y$  in strict lambda notation becomes  $\lambda x.\lambda y.x + y$ .

Applications take abstractions and apply other lambda terms to them using substitution. In a term  $(\lambda x.M)N$ , substitution takes all the instances of the bound variable  $x$  and replaces them with the incoming variable  $N$ , or expressed mathematically as  $(\lambda x.M)N = [N/x]M$ .

With abstractions, applications and variables, any computable function can be expressed. Functions and constants in applied lambda calculus simplify this slightly, allowing the number one to be expressed as the integer 1 rather than the Church Numeral encoding  $\lambda f.\lambda x.fx$  in lambda calculus [32].

## 2.1.2 Alpha Conversion

Complexity is increased when substituting variables into abstractions via applications. The lambda term  $\lambda x.\lambda y.y * x$  multiplies the first input to the function with the second, with  $x$  and  $y$  as distinct, different terms. Applying  $y$  to this function would result in the application  $(\lambda x.\lambda y.y * x)y$  which reduces to  $\lambda y.y * y$ , the mathematical representation of multiplying the input by itself. This is incorrect, the meaning of the function has been changed because of the letter chosen to act as the input, and is an example of variable capture.

Alpha equivalence is the notion in lambda calculus that the choice of variable names does not matter,  $\lambda y.y$  is alpha-equivalent to  $\lambda a.a$ . This notion is used to prevent variable capture through a process called alpha conversion, the process of comparing the incoming variable to the terms inside the abstraction and changing variable names to eliminate capture.

De Bruijn notation is an alternative method of handling this issue, eliminating the need for variables altogether and instead expressing abstraction terms as an index representing which abstraction scope they refer to [16].

The Barendregt Convention assumes bound variables are suitably chosen to avoid capture, a convention which is taught in the Theory of Computation course. From this and alpha conversion, a set of rules for substitution can be created which rename terms to avoid variable capture as substitution is taking place, defined as follows [1]:

$$[t/x]y = \begin{cases} t & \text{if } y = x \\ y & \text{if } y \neq x \end{cases} \quad (2.1)$$

$$[t/x](t_1 t_2) = [t/x]t_1 [t/x]t_2 \quad (2.2)$$

$$[t'/x](\lambda y.t) = \begin{cases} \lambda y.t & \text{if } x = y \\ \lambda z.[t'/x][z/y]t & \text{if } x \neq y \wedge z \notin FV(t) \cup FV(t') \end{cases} \quad (2.3)$$

## 2.1.3 Reduction

Reduction is a formal method for evaluating lambda terms, of which there are different levels. Alpha conversion (or alpha reduction) [14] is the process of renaming variables to avoid variable capture. Beta reduction focuses on finding the result of applications, discussed in detail below. Delta reduction applies functions to constants arithmetically. Finally, eta reduction takes the output of delta- or beta-reduction and removes any leftover abstraction term in an attempt to simplify the final result [29]. Not all of these reduction steps are necessary, but each simplifies the term in some way, terminating with a final result.

### Beta Reduction

Following alpha conversion, beta reduction is ready to be performed abiding by the Barendregt Convention. A definition is first needed which states when a lambda term has reached its simplest form and can be evaluated no further. This is known as the Normal Form of a lambda term [2], and is defined as a term with no beta-redex [15] or which contains no subterms of the form  $(\lambda x.M)N$  [12]. As previously stated, beta reduction is the process of

finding the result from the application of a function, and can be performed by following a number of rules [8].

1.  $(\lambda x.M)N = [N/x]M$   
 $(\lambda x.M)N$  is equal to  $M$  with all instances of  $x$  replaced by  $N$  (substitution)
2.  $M = M$   
A term is equal to itself
3.  $\frac{M=N}{N=M}$   
If  $M = N$  then  $N = M$  through symmetric equality
4.  $\frac{M=N \quad N=L}{M=L}$   
If  $M = N$  and  $N = L$  then  $M = L$  through transitivity
5.  $\frac{M=N}{MZ=NZ}$  and  $\frac{M=N}{ZM=ZN}$   
If  $M = N$ , then  $Z$  applied to  $M$  is the same as  $Z$  applied to  $N$ . Similarly if  $M = N$  then  $M$  applied to  $Z$  is the same as  $N$  applied to  $Z$
6.  $\frac{M=N}{\lambda x.M=\lambda x.N}$   
If  $M = N$ , then  $(\lambda x.M) = (\lambda x.N)$

Different choices made in the application of these rules can generate different results. The Church-Rosser theorem is a fundamental theorem in the subject of lambda calculus, and leads to a useful axiom: a term in the lambda calculus has at most one normal form [6] meaning different reduction strategies cannot produce different normal forms. However, if a lambda term does have a normal form, not all methods of reduction reach it [14].

The Theory of Computation course teaches two approaches to beta-reduction, call-by-value and call-by-need. Call-by-value is a leftmost innermost strategy, [8], meaning in a term  $MN$ ,  $M$  gets reduced to its normal form before  $N$  (the left term is reduced first). In a term  $(\lambda x.M)N$ ,  $N$  is evaluated before it gets substituted into  $(\lambda x.M)$ . For this reason, call-by-value is sometimes referred to as applicative order reduction, as the arguments are reduced before they are applied [30].

Call-by-need is similar but with one key difference, rather than reducing  $N$  and then substituting the result into  $M$ , the original  $N$  is substituted into  $M$  before  $[N/x]M$  is reduced together as a whole. This is a subtle but very important difference, call-by-need can be proven to always turn a lambda term into its normal form if one exists [14], unlike call-by-value. Because of this property, call-by-need is often referred to as normal-order-reduction.

## 2.1.4 Typing Rules

A desirable property in software engineering is static checking at compile time. This can be done with lambda calculus by introducing types, which can be used to verify that a lambda term is performing valid computations before being evaluated arithmetically. For example, the lambda term  $(\lambda x.x+1)y$  is a valid term, but is not runnable if  $y$  is a boolean value such as `TRUE` or `FALSE`.

To use this, a type definition is first needed. A type  $T$  can be of ground type  $A$  (boolean or integer), or function type  $T \rightarrow T$  (taking a type  $T$ , and converting it into another type  $T$  through a function) [8]. In order to type-check the validity of a lambda term, a typing context  $\Gamma$  is expressed which declares a set of variables to be associated with specific types. From this, three typing rules can be expressed [8]:

- For variables, if  $x$  of type  $T$  belongs to the typing context  $\Gamma$ , in  $\Gamma$ ,  $x$  is of type  $T$ :

$$\frac{x : T \in \Gamma}{\Gamma \vdash x : T} TVar \quad (2.4)$$

- For abstractions, if in a typing context  $\Gamma$  where  $x$  has type  $T$ , if  $M$  has type  $U$ , then  $\lambda x.M$  has type  $T \rightarrow U$ :

$$\frac{\Gamma, x : T \vdash M : U}{\Gamma \vdash \lambda x : T. M : T \rightarrow U} TAbs \quad (2.5)$$

- In an application  $MN$ , if  $M$  has type  $T \rightarrow U$  and  $N$  has type  $T$ , then  $MN$  has type  $U$

$$\frac{\Gamma \vdash M : T \rightarrow U \quad \Gamma \vdash N : T}{\Gamma \vdash MN : U} TApp \quad (2.6)$$

This is all that is needed to statically analyse a lambda term, and determine whether or not a lambda term is typable, or feasible when considering typing and typing rules. The rules can be used to determine what type the lambda term should return, provided information about the types of some of the variables.

## Type Inference

In a lambda expression, the type of a term can often be inferred without any explicit type declaration. For example, in an expression  $\lambda x.x + 1$ ,  $x$  can be assumed to have type `int`, as  $+$  is an integer operator. This is type inference, and is the process of giving a variable a generic type (or supertype) and then refining this type based on gathered information until a minimal type is reached. [8] [3] [34]

This gathering of information can be split into two categories, structural and syntax [3]. Structural is done using lambda typing rules, for example using the fact that abstractions are function types of form  $T \rightarrow T$  to find the final type of the term  $\lambda x.x + 5$ . The syntax of the term can also be used to gather information, for example using arithmetic or boolean operators to determine the type of a variable used in that expression.

# Chapter 3: Analysis

As defined in the Introduction of this report, the focus of this tool is to help students taking the Theory of Computation course gain a better grasp lambda calculus. Because of this, the Theory of Computation course material will be used to help determine the objectives of the tool in the following sections.



## 3.1 Lambda Calculus

The Theory of Computation course details [11] state two key intended learning outcomes relating specifically to Lambda Calculus: evaluate expressions in lambda calculus according to the definition of reduction relation, and determine whether or not expressions in lambda calculus are typable in the simple type system. From this, the following objectives can be selected:

- The Barendregt Convention is being chosen as a means of avoiding variable capture over De Bruijn notation, since it aligns more closely with what is taught on the course and therefore will better support student learning
- Simply typed applied lambda calculus will be implemented
- Beta reduction will be included, with options to evaluate terms using both call-by-value and call-by-need reduction strategies as per what is taught
- Typing derivations will be implemented in order to type check the lambda term and evaluate the final type of the reduced function

Further to what is taught in the course, there are other options which could be included in such an interface. Delta reduction has been discussed throughout the background section but is not taught explicitly, however lecture slides do include constants being added together arithmetically, performing delta reduction implicitly. Because of this, the resultant lambda term will have arithmetic constants evaluated but delta reduction will not be explicit. Eta reduction has also been discussed, but is not an integral step to lambda calculus, and most resources teaching lambda calculus miss it out [1] [2] [8] [14] [26] [12]. Since manipulating the final result in a way students have not been taught could cause further confusion, eta reduction will not be included in the final program.

## 3.2 Abstract Syntax

A way of representing the structure of a lambda term is needed too take a term and turn it into a program which can evaluate lambda terms. From this, the resultant program can use the defined rules to evaluate terms.

The idea behind this is to turn a sequence of character strings (the surface syntax of an expression) into an abstract syntax (a representation of the meaning of the expression). A lexer is used to do this, which converts a character sequence into a series of identifiers called tokens, as defined by the grammar of a language (in this case lambda calculus). A parser is then used which turns this sequence of tokens into a tree known as an Abstract Syntax Tree, which can then be traversed and explored to create compilers and interpreters of a language [25]. The applications of this are wide, it can be used to evaluate arithmetic expressions (Figure 3.1), in syntactic analysis used for natural language processing [4] (Figure 3.2), or to evaluate lambda calculus terms as done in this project (Figure 3.3).

### 3.2.1 ANTLR

One such tool which can help represent the structure of a term based on a particular grammar is ANTLR. ANTLR is a parsing tool which is used to build and explore syntax

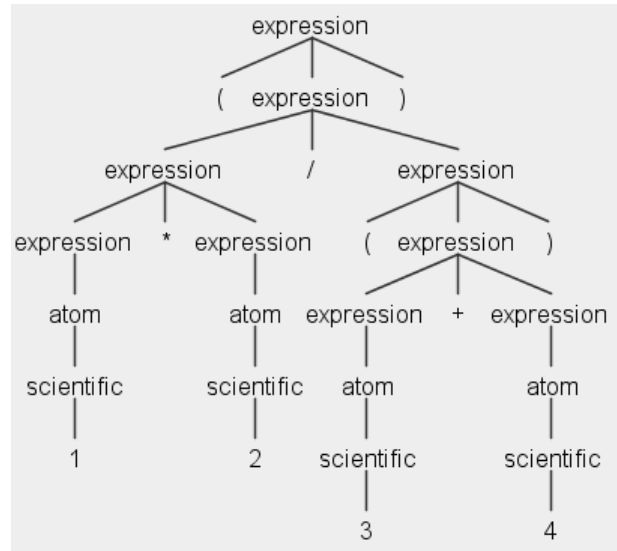


Figure 3.1: Syntax Tree of an Arithmetic Term (generated using code from [9])

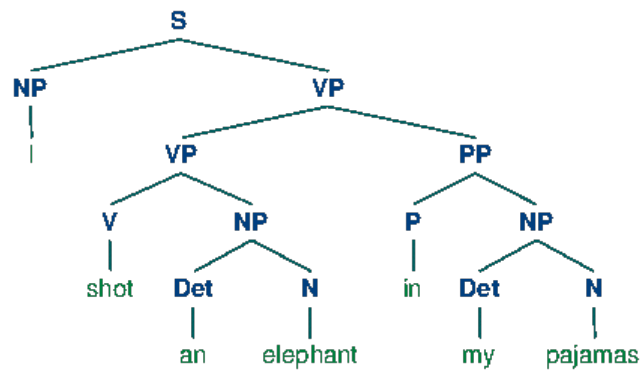


Figure 3.2: Syntax Tree for Natural Language Processing [4]

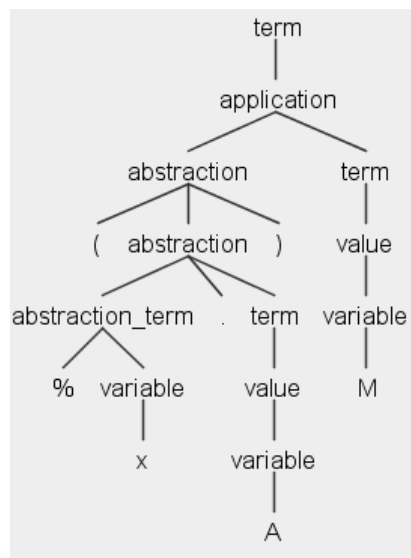


Figure 3.3: Lambda Calculus term  $(\lambda x.A)M$  abstract syntax tree

trees from character strings. It has a lexer to generate tokens, and has a parser to turn this stream of tokens into an abstract syntax tree which can then be evaluated using either a listener or a visitor [23]. It has been used in a wide variety of contexts, including Twitter (for query parsing) and the NetBeans IDE (for parsing C++) [23].

### 3.3 Interface

Once a program has been created which can parse and evaluate lambda terms, an interface is needed to allow this to be used by students. The interface should be simple to understand, should include a place for user input which will allow students to input a lambda expression and then see the resultant reduced expression. It should also include an input which will allow the user to select different types of beta-reduction, and an output which will show the typability and evaluated type of the returned expression.

## Chapter 4: Aims and Objectives

From the background and analysis discussed above, a set of key aims and objectives can be defined. These eight objectives have been defined below, and have been labelled using the MoSCoW system, (must-have, should-have, could-have and would-like-to-have) [7]. The program will include:

- Must Have: The ability to evaluate the validity of lambda calculus terms, so students can enter terms and check whether or not they are valid. This will be done by turning the user input lambda term into an Abstract Syntax Tree using ANTLR
- Must have: Applied lambda calculus, to allow the input of expressions
- Must have: The ability to alpha convert these terms, to help students understand alpha conversion and equivalence. This will be done by navigating the abstract syntax tree generated by ANTLR
- Must have: A web interface as the front-end for this program to allow users to use the tool easily
- Should have: Beta reduction used to evaluate terms to normal form to help students understand beta reduction. This will also be done by navigating the abstract syntax tree generated by ANTLR
- Should have: Types are included, to help students understand types and typing rules
- Could have: Information is given to the user about the typing context from which typing validity has been determined
- Could have: Different reduction strategies being offered, to give students the option to understand call by value/call by need evaluation
- Would like to have: Type inference performed to determine the output type of a lambda term
- Would like to have: Arithmetically evaluate these terms using delta reduction

- Would like to have: Details given to the user on how the term was evaluated using beta-reduction

# Chapter 5: Design and Implementation

## 5.1 Design

The overall program is split into a number of distinct elements. Firstly, a grammar has been written which defines the syntax of lambda terms and allows a term to be broken down into individual tokens. ANTLR uses this grammar to process lambda terms, creating an Abstract Syntax Tree for each incoming expression [23], a process which has been well documented and therefore was not deviated from. A tree traversal mechanism forms the main section of code which navigates the abstract syntax tree and performs operations to determine the resultant expression. This is the section of code which performs the evaluation, and therefore needed to be designed accurately with the rules of beta reduction and lambda calculus in mind. A web interface has been used to house this program, giving the user a simple and convenient way to interact with the underlying system.

The relevant sections have been discussed below, detailing design and key decisions which needed to be made before implementation began.

### 5.1.1 Grammar

The key aim in creating this grammar was creating a syntax which sticks as closely as possible to the rules of lambda calculus. Lambda Calculus grammar is already clearly defined, with a lambda term being either a variable, an abstraction or an application [12]. Applied lambda calculus adds functions and constants to this definition [28], resulting in the following lambda grammar. A lambda term is one of the following:

- | An application (of form `[term] [term]`)
- | An abstraction (of form `[abstraction_term].[term]` where `[abstraction_term]` is defined by  $\lambda[\text{variable}]$ )
- | A function (of form `[term] [operation] [term]`)
- | A value (of form `[variable]` (the letters a-z) or `[number]` (constant))

With the addition of types, the grammar adds the option of a `:[type]` term to each variable, with each type being either a ground type (bool, int or none), or in the form `[type]  $\rightarrow$  [type]`. This follows the standard syntax for typing used in the lecture material [12] [8], and consistency with this syntax allows students on the Theory of Computation course to input a lambda term directly from the lecture slides with minimal adjustment.

### 5.1.2 Expression Evaluator

The fundamental component of this evaluator is a tree traversal, which navigates through the token nodes and performs different operations depending on the type of node encountered. Abstractions take an incoming value if one exists, and substitute it into the body of

the abstraction. This allows an evaluated expression to be built up, and a result determined.

ANTLR provides two mechanisms for traversing an abstract syntax tree: listeners and visitors. A listener is a passive way of evaluating a syntax tree, used within an ANTLR Walker class which traverses the tree using a depth-first approach and triggering methods from the listener as it enters and exits each token [23]. These listener methods are unable to return values, so expressions and evaluations have to be handled using separate objects within the listener class. As the walker traverses the tree, the listener builds up a running evaluation of the term, returning the result when it exits the topmost node [31].

Unlike listeners, visitors control their own traversal of the tree. By visiting the children of each node encountered explicitly, the path they take around the tree can be controlled [23]. Visitors also allow custom return types, meaning nodes can return their resultant expressions directly to their parent node and do not have to rely on separate objects [31].

With beta reduction, different methods take different approaches to evaluating terms. In an application MN using a call-by-value approach, N is evaluated before M. In a call-by-need approach, N is passed into M before being evaluated. This means that depending on the type of reduction selected, the evaluator will have to traverse the tree in a different order, suggesting visitor over listener is more appropriate for this task. Furthermore, visitor methods being able to return values directly to their parents as opposed to relying on separate objects will support evaluation happening as the tree is being visited, due to the large amount of data being passed around the tree.

The visitor therefore was the main code written in this project, with an ANTLR generated parser is passed to a custom visitor interface. Different visitors are defined for each of the beta-reduction strategies being implemented, since each method takes a different approach to evaluating expressions.

The visitor returns three separate items upon returning to the topmost node: the evaluated value of the input lambda term, whether or not the term is typable, and the type of the final expression. It also returns details of any errors where applicable, for example syntax errors which ANTLR is unable to parse, or cases where the normal form of a lambda term does not exist.

### 5.1.3 Web Interface

The web interface allows the user to input a lambda term along with the type of any variable. It also allows the user to select which reduction strategy they would like to have the term evaluated by, with call-by-need being selected as the default.

The code running the interface and the underlying evaluation code are kept as separate entities, communicating through input parameters and return statements. This is to ensure the code is kept as modular as possible, allowing the evaluator code to be run using multiple different mechanisms, for example through a web interface and a command line for testing purposes. Any data the user enters is passed to the underlying code, which processes the term and return any output via a return statement, being read and displayed to the user.

## 5.2 Implementation

Despite the examples in the ANTLR Reference Guide [23] all being written in Java, Python was chosen as the main language for this project. ANTLR offers support for Java and Python, both of which are regularly listed in the top three programming languages being used in industry from various sources [13][24][5]. Since the MSc Software Development course which this dissertation is being written for teaches only Java, Python was chosen in order to expand the number of programming languages encountered throughout the course. Therefore the majority of the code for this project will be written in Python, with HTML being used for the web interface.

### 5.2.1 Grammar

ANTLR requires that the grammar of a language is contained in a .g4 file which defines the parser and lexer rules for the language. A section of the grammar written for lambda calculus is shown in Figure 5.1, and indicates the parser rules for some of the token nodes, alongside the lexer rules which are shown in Figure 5.2. This grammar is consistent with the syntax defined in the background section of this report, with slight modifications such as the addition of parentheses and an explicit definition for boolean and integer values.

### 5.2.2 Abstract Syntax Tree

Having defined a grammar, ANTLR can then be used to create an abstract syntax tree, which can be seen for an example term in Figure 5.3. This is very simple to do using ANTLR libraries, and involves passing the input term through a number of ANTLR generated classes before passing the result through a parser which turns the tokens into an abstract syntax tree [33]. The lexer and parser are created by ANTLR when the grammar file is compiled, resulting in a syntax tree which processes terms based only on the syntax of the defined language.

The process of creating a tree from the grammar and an incoming term is well documented, and so no modifications need to be made, with the exception of attaching a custom Error Listener to both the lexer and parser as expanded on later on in this Implementation section.

### 5.2.3 Expression Evaluator

This is the largest block of code in the program, and defines how the incoming lambda term should be evaluated depending on the input given by the user. For the purposes of simplicity this code can be split into three distinct sections, alpha conversion, beta reduction and typing rules.

#### Alpha Conversion

The alpha conversion is based on the rules for explicit alpha conversion with substitution, as defined by [1], and as discussed in detail in a previous chapter. The final rule is the rule to be focused on and so will be repeated here for clarity:

```

/* Parser Rules */
term
: abstraction
| function
| value
| application
;

application
: application term
| abstraction term
| value term
| function term
| LBRACKET application RBRACKET
;

abstraction
: abstraction_term '.' term
| LBRACKET abstraction RBRACKET
;

abstraction_term
: '%' variable
;

```

Figure 5.1: Parser Rules

```

/* Lexer rules */
NUMBER : [0-9]+ ;
BOOL : 'TRUE' | 'true' | 'True' | 'FALSE' | 'false' | 'False' ;
VARIABLE : [a-zA-Z] ;
ADD : '+' ;
SUBTRACT : '-' ;
MULTIPLY : '*' ;
DIVIDE : '/' ;
POWER : '^' ;
LBRACKET : '(' ;
RBRACKET : ')' ;
AND : '&' ;
OR : '|' ;
GT : '>' ;
LT : '<' ;
EQ : '==' ;

WS : [ \t\r\n]+ -> skip ;

```

Figure 5.2: Lexer Rules

$$[t'/x](\lambda y.t) = \begin{cases} \lambda y.t & \text{if } x = y \\ \lambda z.[t'/x][z/y]t & \text{if } x \neq y \wedge z \notin FV(t) \cup FV(t') \end{cases}$$

This defines the process of finding a variable  $z$  that does not appear in the free variables of the incoming term or the existing term, replacing all bound variables  $y$  with this new term  $z$ , and then substituting in  $t'$  as normal.

The alpha conversion code written for this project follows this process. First, the set of free variables in the term are determined, by taking the set of all alphabetic characters in the term and eliminating the bound variables. These variables are then replaced with letters that are not in the list of free variables, in situations where a clash in free variables between the two expressions are found. This produces an alpha-converted term, into which substitution can be performed using a simple `String.replace()` method in Python.

## Beta Reduction

Two beta reduction strategies are taught on the Theory of Computation course, call-by-value and call-by-name. The differences between these two strategies have been discussed in detail, but the key difference is in the order in which substitution and evaluation occurs in an application  $MN$ .

Because of these differences in evaluation strategy, two separate visitors are used to limit the use of control coupling [19]. Since they share a lot of functionality, a `BaseVisitor` was defined which contains all common code between these two strategies. The two call-by-value and call-by-need visitors are subclassed from this base visitor, and define their unique behaviour for the application and abstraction terms. Both visitors pass information to the web interface about the steps being taken to evaluate the term as the tree is navigated. This output is given to the user alongside the evaluation result through a clickable link which opens a separate webpage containing this information. This can be seen in Figure 5.7.

The abstraction term differs between the two methods due only to typing, since determining the type of a term happens during the evaluation of that term. In call-by-value, the type of  $N$  is known before substitution, so can be carried throughout the function. In call-by-need, the whole term needs to be type checked after substitution has happened to determine the type of  $M$  with  $N$  incorporated. Two separate abstraction methods were therefore needed, with common code moved into a method defined within the base visitor.

## Typing Rules

Lambda Calculus has clearly defined typing rules which are taught in the Theory of Computation course. These rules have been discussed in more detail previously. Each type can either be of ground type  $T$ , or function type  $T \rightarrow U$  under a typing context  $\Gamma$  [12].

Since the `visitNode` methods already evaluate their children nodes, these typing rules can be integrated directly into these nodes, with each `visitNode` method returning a value and a type. When a type is determined for a variable, this is added to a global typing context associated with the term. This is given to the user via a separate webpage linked to the output type. Once the term is evaluated, shown in Figure 5.7. Type checking happens



throughout the code in each method where evaluation occurs, and is implemented as follows:

#### *Variable*

Variables simply return the type given to them by the user, as defined by Rule 2.4. Any number is given type `int`, and any boolean values `TRUE` or `FALSE` are given type `bool`.

#### *Abstraction*

Users enter types in the form `x:type`, which can be applied to any variable included in the term. The abstraction method takes the input type `T`, and joins it with the output type with `->`, to become of function type `input->output`. The input and output types are either specified by the user, or determined through type inference.

#### *Application*

For an application `MN`, application typing involves iteratively removing the first ground type of `M` and the first ground type of `N` until the type of `N` is `None`. For example, if `M` is of type `int->bool->int` and `N` is of type `int->bool` the typing of application `MN` is of type `int` through the following process:

$$\begin{array}{ll} M : int \rightarrow bool \rightarrow int & N : int \rightarrow bool \\ M : bool \rightarrow int & N : bool \\ M : int & N : None \end{array}$$

If at any point the first ground type of `N` does not match the first ground type of `M`, the typing is invalid (for example if `N` was of type `bool->bool` in this case). This is implemented using string manipulation, splitting a type `T` into its individual ground type using the `split()` method in Python, then comparing and removing the types of `M` and `N` iteratively.

Type inference uses the typing rules defined above to determine the type of an output term where possible. Abstractions are initially given function type `None->None`, and variables are given ground type `None`. This is inference using structural information.

In the `visitFunction` node, syntactical information can be used to determine the input and output types of terms by examining the operation term. The following is declared in code:

- The operations `{&,|}` take two boolean values and return a boolean. Therefore the type of the incoming and outgoing term can be inferred to be of type boolean
- The operations `{+,-,*}` take two integer values and return an integer. Similar to above, the incoming and outgoing term can be inferred to be of type integer
- The operations `{==,>,<}` take two integer values and return a boolean, it can therefore be inferred that the input term is of type integer, and the output is of type boolean.

The output type of a function is used to determine the output type of its parent node. In an abstraction, this type inference can be used to determine the type of its input term, for example the lambda term `λx.x + 1` can be inferred to have type `int->int` despite no

input type given by the user.

There is a limit to this type inference, the function will infer the input type when there is only one operation term. For example, the term  $\lambda x.(x == 1) \& b$  contains two operations, `==` and `&`. In this case, working out the input type is more complicated, since the typing has to be broken down into sub-functions, which need to determine what they think the input type should be, which then needs to be examined collectively as a complete term. The type of this output is returned by the program therefore as `None->bool`. This is correct, just not as minimal as its alternative type `int->bool`.

While narrowing type inference to its most minimal state is definitely possible, it is not the key goal in the project, and is a small edge-case when it comes to improving student understanding. Because of this, and due to the finite nature of this project, this limitation has been deemed acceptable, with other tasks which more greatly contribute to students understanding of lambda calculus being prioritised.

## Error Handling

Each visit method in the visitor returns an evaluated result and a type, which is used to build the evaluated term and determine the final result, type and type validity of the input expression. However, during this process, there are a number of occurrences which could stop the program from being able to return a final result. These are broken down into two key issues: syntax errors and occurrences where a term doesn't have a normal form (and therefore cannot be evaluated to termination). These have been summarised below.

### *Syntax Errors*

Lambda terms often include nested parentheses. It is easy for a user to mismatch brackets, such as forgetting to close an open bracket. Because of this, before the term is passed to ANTLR, the code checks to ensure brackets are matched correctly, and if not, returns an error to the user informing them of the issue and asking them to re-enter.

Aside from this, syntax errors specific to lambda calculus are likely to occur which cannot be picked up until the term is passed through a lexer. To handle this, ANTLR's `ErrorListener` class is overridden, instead throwing a custom `SyntaxTokenError` exception which is caught by the program and passed back to the user.

### *Recursion Errors*

Recursion errors are thrown in Python when a maximum stack depth is reached in order to protect Python from crashing [10]. This is the exception which is thrown when a lambda term has no normal form, as the code keeps trying to evaluate the term before eventually throwing a `RecursionError`.

To handle this, the recursion limit for the code is set to 200, smaller than the python default to prevent excessive time waiting for the program to fail, but large enough to ensure that any reasonable lambda term can be processed. A recursion limit for the code equates to a lambda term with approximately 50 nested abstractions. While this is a limitation, it's deemed to be a reasonable one, as limiting students to 50 nested abstractions as opposed to 100 is unlikely to impact student understanding. If a recursion error is thrown, it is caught by the code in a `try/except` block, and an error message informing the user that a normal form cannot be found for this term is given.

### 5.2.4 Web Interface

Flask is a web framework designed for use with Python, and allows python scripts to be controlled from a web interface [21]. The websites' front-end is written in HTML, and is connected to the python file using Flask's `render_template()` function.

Users enter lambda terms in the input dialogue box, entering a % sign as a lambda symbol. Using a HTML onkeydown event, the interface automatically changes this symbol to a  $\lambda$ , allowing users to easily enter lambda terms. The complete web interface can be seen in Figure 5.4, Figure 5.5 and Figure 5.6.

Clicking the *Check Expression* button on the web interface sends a HTTP post request to the Flask server, which takes the value held in the user input box along with the selected reduction strategy, and sends this to the main lambda program. The output is returned to the HTML code through Flask, and the webpage is re-rendered to display the results.

### 5.2.5 Complete Program

The complete class diagram for the final program can be seen in Figure 5.8, and shows how the web interface connects to the main lambda code using the `route.py` file which integrates Flask. It also shows how the ANTLR generated classes are utilised, and how each evaluation method is implemented using a separate visitor.

## Chapter 6: Testing

The aim of this project is to build a web application which will verify and evaluate lambda calculus terms input by the user, with the goal of improving student understanding. This can be broken down into two key goals for the created tool, the interface should evaluate lambda terms accurately and should be useful for students learning functional programming.

Because of this, the testing will be broken into sections in accordance with these two goals. The first set of tests will evaluate the accuracy of the lambda term evaluation. The second set of tests will test the usability of the interface and whether or not it meets the goal of improving student understanding, by performing a user test with students who have an understanding of computer science and performing statistical analysis on the results.

### 6.1 Accuracy of Lambda Term Evaluation

In order to test the accuracy of the evaluation, the tests were split into three categories: alpha-conversion, beta-reduction and typing rules. For each category, key points of potential failure were identified. Unit tests written using the pytest framework were used to test the beta-reduction and typing rules, split into 13 different test modules. The alpha-conversion was not tested for using unit tests due to alpha equivalence being difficult to unit test for, since  $\lambda x.x$  and  $\lambda a.a$  are equivalent terms but not string equivalent.

The results of the tests are shown in Figure 6.1 and 6.2 below, with the unit test results shown in Figure 6.3 and 6.4. Each table includes the input to test, the motivation for testing

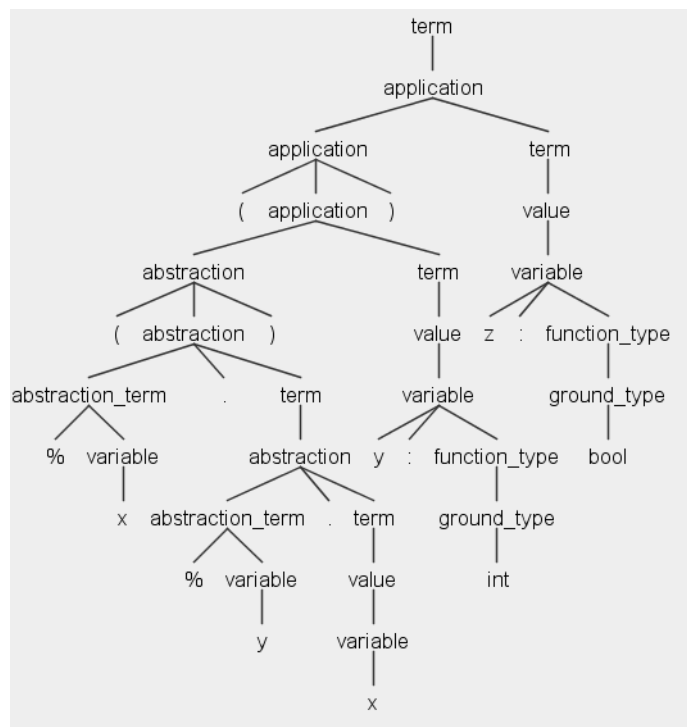


Figure 5.3: Abstract Syntax Tree Example for  $((\lambda x.\lambda y.x)y:\text{int})z:\text{bool}$

## Lambda Calculus Expression Evaluator

Enter the term in the box below, the result will tell you the output, whether it is typable and what type it should be

To enter  $\lambda$ , type %

Types can also be entered for any variable, in the form  $x:\text{type}$  (*bool, int and none types allowed*)

Enter expression here

**Evaluation Method:** ☐ Call By Value (applicative order) ☒ Call By Name (normal order) ☐ Alpha Conversion Only

Check Expression

Note: this program only allows single input abstractions. To enter a multi-input abstractions, nest them. So  $\lambda xy.xy$  becomes  $\lambda x.\lambda y.xy$

Need more information? Have a look at the [User Guide](#) for a list of commands and operations

Spot an issue? Code can be found [here](#)

Figure 5.4: Web Interface

## Lambda Calculus Expression Evaluator

Enter the term in the box below, the result will tell you the output, whether it is typable and what type it should be

To enter  $\lambda$ , type %

Types can also be entered for any variable, in the form x:type (*bool, int and none types allowed*)

Enter expression here

**Evaluation Method:** ☐ Call By Value (applicative order) ☒ Call By Name (normal order) ☐ Alpha Conversion Only

Check Expression

Input =  $(\lambda a.a + (\lambda x.x + 1)3)4$

Result =  $(4 + (3 + 1)) = 8$  by arithmetic reduction ([click here for evaluation details](#))

Valid typing = True under typing context ([click here](#))

Type returned = int

Note: this program only allows single input abstractions. To enter a multi-input abstractions, nest them. So  $\lambda xy.xy$  becomes  $\lambda x.\lambda y.xy$

Need more information? Have a look at the [User Guide](#) for a list of commands and operations

Spot an issue? Code can be found [here](#)

Figure 5.5: Web Interface With Returned Result

## Lambda Calculus Expression Evaluator

Enter the term in the box below, the result will tell you the output, whether it is typable and what type it should be

To enter  $\lambda$ , type %

Types can also be entered for any variable, in the form x:type (*bool, int and none types allowed*)

Enter expression here

**Evaluation Method:** ☐ Call By Value (applicative order) ☒ Call By Name (normal order) ☐ Alpha Conversion Only

Check Expression

Input =  $(\lambda x.\lambda y.y)((\lambda z.zz)(\lambda z.zz))$

Normal form cannot be found - does this term have a normal form?

Note: this program only allows single input abstractions. To enter a multi-input abstractions, nest them. So  $\lambda xy.xy$  becomes  $\lambda x.\lambda y.xy$

Need more information? Have a look at the [User Guide](#) for a list of commands and operations

Spot an issue? Code can be found [here](#)

Figure 5.6: Web Interface With Error

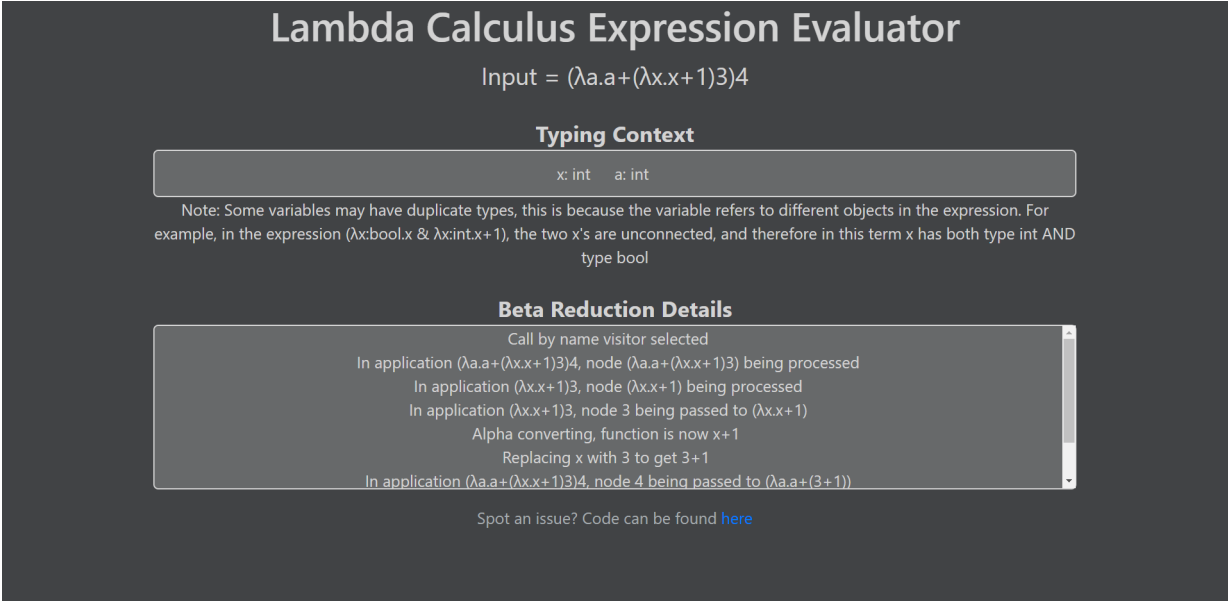


Figure 5.7: Web Interface

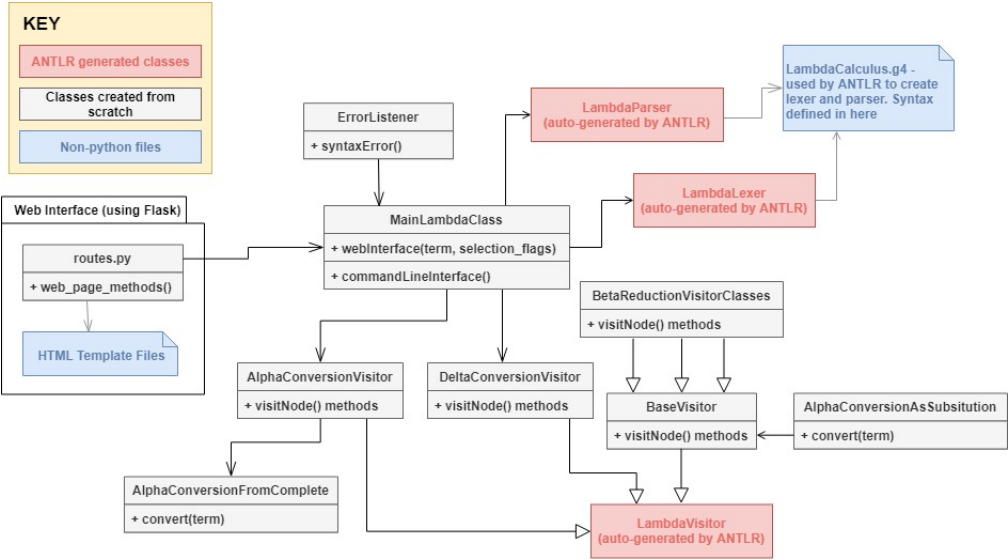


Figure 5.8: Complete Code Class Diagram

that input and the results, with the name of the unit test where applicable. The unit test outputs have been split into those which have passed and those which have failed for clarity of reading.

Figures 6.1 and 6.2 demonstrate that in the majority of cases which have been tested, the evaluation is accurate and does yield expected results. Any instances where the output is not as expected have been discussed in the Evaluation section of this report.

## 6.2 Interface Usefulness and Usability

In order to test the usability and usefulness of the interface, a user test was conducted using participants who had a computer science background, to simulate those on the Theory of Computation course. Statistical analysis using quantitative data has been used to determine whether or not the interface created helps to improve understanding of lambda calculus. Qualitative feedback given by participants in the user test is also used to support any conclusions drawn and a discussion on user ideas and possible improvements has been made.

### 6.2.1 Testing

In the user test, participants were first given a Participant Information sheet and to sign a Participant Consent Form, which stated they were happy with the results of their user test to be included in this dissertation, provided their answers will be kept anonymous as will be done here. They were then asked to read a document giving them a brief introduction to lambda calculus, which was written for this test and was based on the Theory of Computation lecture slides [8]. This was done to simulate students who have been introduced to lambda calculus, but are still very new to the subject.

Having read a brief summary, participants were asked to rate their understanding of lambda calculus on a scale of 1-10, before being given the interface to test. They were provided with a few sample lambda terms to test if they wanted to, but were not obligated to use them. This was to simulate students who may be using the tool to evaluate a particular lambda expression they have found in the lecture slides or online.

Finally, users were given a questionnaire which again asked them to rate their understanding of lambda calculus having used the interface for a short period of time, so their understanding could be compared before and after using the tool. The questionnaire also asked a number of qualitative questions, asking participants to give feedback on whether or not they would use this tool if they were studying lambda calculus and what they thought of the interface, along with any suggestions they had for possible improvements.

### 6.2.2 Results

Although the test was not conducted with a time limit, each participant spent approximately five minutes using the interface before stating they were ready to move on. As can be seen from Figure 6.5, in the five minutes participants interacted with the interface their understanding improved by 40% on average, with users who started with a lower understanding of lambda calculus seeing the most improvement. 100% of participants said they

Input	Motivation	Expected	Output	Result
<i>Alpha Conversion</i>				
$(\lambda x.x*y)y$	Variable capture protection test	$(\lambda x.a*x)y$ or alpha-equivalent	$(\lambda x.x*a)y$	✓
$(\lambda x.\lambda y.yx)y$	Multi-input, risk of nested variable capture	$(\lambda x.\lambda z.zx)y$ or alpha-equivalent	$(\lambda x.\lambda b.bx)y$	✓
<i>Beta Reduction</i>				
$(\lambda x.\lambda y.x)y$	Alpha conversion simultaneous to substitution	$\lambda z.y$ or alpha-equivalent Type Validity = True Type = None $\rightarrow$ None	$(\lambda a.y)$ Type Validity = True Type = None $\rightarrow$ None	✓

Figure 6.1: Evaluating the Accuracy of Lambda Term Evaluation - Alpha Testing

Unit Test Name	Input	Motivation	Expected	Result
<i>Beta Reduction</i>				
<i>Call-by-value</i>				
test_call_by_value()	$(\lambda x.\lambda y.y)((\lambda z.zz)(\lambda z.zz))$	Testing difference between call-by-value and call-by-need	No normal form	PASS
<i>Call-by-need</i>				
test_call_by_need()	$(\lambda x.\lambda y.y)((\lambda z.zz)(\lambda z.zz))$	Testing difference between call-by-value and call-by-need	$\lambda y.y$ Type Validity = True Type = None $\rightarrow$ None	PASS
<i>Both</i>				
test_multi_input()	$(\lambda b.b+(\lambda a.a+b)3)4$	Multi-input	$4+3+4 = 11$ Type Validity = True Type = int	PASS
test_scope_of_bound_variables()	$(\lambda x.x+\lambda x.x)3$	Scope of bound variables	$3+\lambda x.x$ Type Validity = True	PASS
test_order_of_operations()	$(\lambda x.x*5)3(\lambda y.y+2)$	Order of operations, applications are left associative so $(\lambda y.y+2)$ should stay in the result	$(3*5)(\lambda y.y+2)$ Type Validity = True	PASS
test_int_to_bool()	$(\lambda x.x>4)3$	Testing int $\rightarrow$ bool functions	False Type Validity = True Type = bool	PASS
<i>Typing Rules</i>				
test_type_clashing()	$(\lambda x:\text{bool}.x+1)$	Type clashing, type of x conflicts with operation type +	$(\lambda x.x+1)$ Type Validity = False	PASS
test_type_valid_clashing()	$(\lambda x:\text{int}.x+1)$	Type validity through bound variable	$(\lambda x.x+1)$ Type Validity = True Type = int $\rightarrow$ int	PASS
test_type_multiple_function_types()	$(\lambda x:\text{bool}.\lambda y:\text{int}.y+1)$	Multi-input valid typing	Type Validity = True Type = bool $\rightarrow$ int $\rightarrow$ int	PASS
test_variable_type_clashes()	$(\lambda x:\text{int}.x:\text{bool})$	Variable type clashes	Type Validity = False	PASS
test_type_inference()	$(\lambda y.y==1)$	Testing type inference	Type Validity = True Type = int $\rightarrow$ bool	PASS
test_multi_operation_type_inference()	$(\lambda x.(x+1)==2)$	Testing multi-operation type inference	Type Validity = True Type = int $\rightarrow$ bool	FAIL
test_order_of_arithmetic()	$(\lambda x.x^2+x^3)$	Testing order of arithmetic operations	12	FAIL

Figure 6.2: Evaluating the Accuracy of Lambda Term Evaluation - Beta and Types Testing



```

test_lambda_evaluator.py::test_call_by_value PASSED
test_lambda_evaluator.py::test_call_by_name PASSED
test_lambda_evaluator.py::test_multi_input PASSED
test_lambda_evaluator.py::test_scope_of_bound_variables PASSED
test_lambda_evaluator.py::test_order_of_operations PASSED
test_lambda_evaluator.py::test_int_to_bool PASSED
test_lambda_evaluator.py::test_type_clashing PASSED
test_lambda_evaluator.py::test_type_valid_clashing PASSED
test_lambda_evaluator.py::test_type_multiple_function_types PASSED
test_lambda_evaluator.py::test_variable_type_clashes PASSED
test_lambda_evaluator.py::test_type_inference PASSED
test_lambda_evaluator.py::test_multi_operation_type_inference SKIPPED
test_lambda_evaluator.py::test_order_of_arithmetic SKIPPED
===== 11 passed, 2 skipped in 4.86 seconds =====

```

Figure 6.3: Evidence of Unit Tests Passing

```

===== FAILURES =====
_____ test_multi_operation_type_inference _____

E       AssertionError: test failed because type validity = true with type = none->bool when I expected int->bool
E       assert ('true' in 'true' and 'int->bool' in 'none->bool')

test_lambda_evaluator.py:261: AssertionError
_____ test_order_of_arithmetic _____

E       AssertionError: test failed because result = 1024 when I expected 12
E       assert '12' == '1024'
E       - 12
E       + 1024

```

Figure 6.4: Unit Tests Failing Output

would use the tool created if they were on the Theory of Computation course, and 100% said they found it useful in improving their understanding.

The percentage of improvement found in Figure 6.5 will now be used to determine whether students are statistically likely to improve in understanding using the tool, and whether or not the improvement is statistically significant or due to chance. This has been done using z-scores and t-testing.

We can prove whether the tool created was successful in improving understanding using z-scores and Statistical Hypothesis Testing. Given an alpha-level of 0.05, this leads to a z-value of 1.645 [18]. With a sample size of 12 and a mean percentage of improvement of 39.1%, we can state a null hypothesis that participant understanding improved less than 25% after using the tool, leading to a z-score of 2.3 which is larger than the z-value of 1.64. This means that the null hypothesis can be confidently rejected, and leads to the conclusion that with a short amount of time using the tool, students will on average likely improve in understanding of lambda calculus by at least 25%.

Calculating the t-value of this data using the equal variance method [17] results in a t-value of 1.93, or approximately 93% confidence that the improvement in understanding is statistically significant. It is acknowledged that the sample size of 12 in this case is small due to time restrictions in the project, and in order for more concrete conclusions to be drawn more participants would be needed, which has been discussed further in the Future Work section of this report. This result is still significant however and supports the hypothesis that the interface created is useful in improving understanding of students learning lambda calculus.

The qualitative questions on the tool revealed that participants found the system simple and easy to use, with some suggestions for improvement. Suggested interface changes included making the colour of the input box darker to differentiate between input boxes and menus, changing the background as some participants found it distracting, and splitting up the output box to make individual elements clearer. Participants also suggested adding additional features such as tooltips to give the user more information on what particular lambda terms meant, and adding a workbook section which would allow users to save the results to a file in order to create a study resource they could go back to later.

Overall the feedback was incredibly positive, with participants commenting that the tool helped improve understanding by adding a practical element to the much more theoretical-based lecture slides. With an original goal for the tool of helping to improve student understanding in lambda calculus, this combined with the Statistical Hypothesis Testing above proves that this goal has been met successfully.

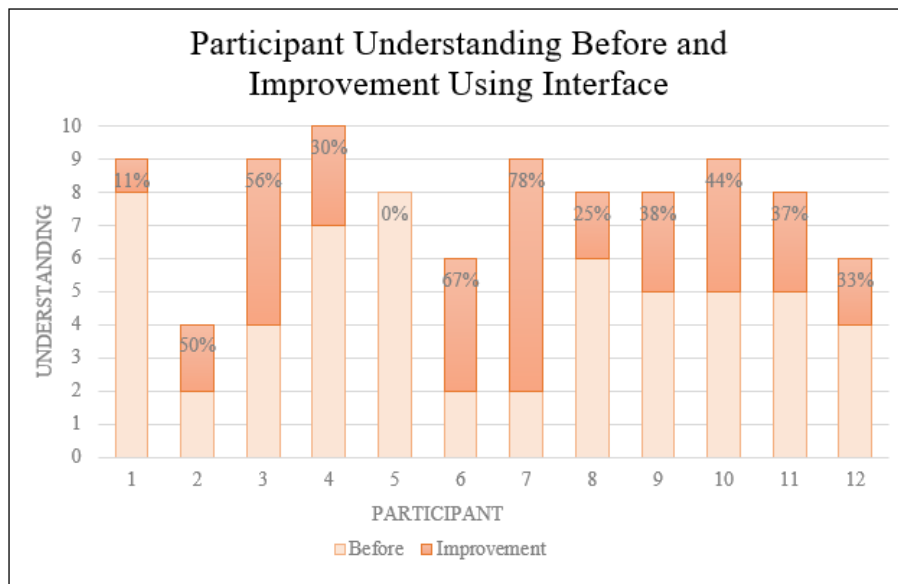


Figure 6.5: Participant Understanding Results

## Chapter 7: Evaluation

As stated in the beginning of this report, the aim of this project is to build a web application which will verify and evaluate lambda calculus terms input by the user. This tool is created with the goal of helping future students learning functional programming to gain a better understanding of lambda calculus.

This chapter will discuss whether or not the project description has been met, and whether or not the tool created could be used as a supplementary resource for students learning lambda calculus on the Theory of Computation course taught at the University of Glasgow.

## 7.1 Accurately Evaluating Lambda Terms

As evidenced by section 6.1, the tool created does accurately evaluate lambda terms, returning the normal form, type and type validity of lambda terms for both call-by-value and call-by-need evaluation strategies. Applied lambda calculus with types is supported, as taught on the Theory of Computation Course.

### 7.1.1 Limitations

Figure 6.2 demonstrates the type inference issue when functions have multiple operations. This has been discussed in detail in the Implementation section of this report, and therefore will not be discussed here. It also shows an issue related to the order of arithmetic operations in a function. Functions in the syntax are set to be right-associative to align with abstractions, which means that the function  $2^2 + 2^3$  is processed as  $2^{(2 + (2^3))} = 1024$  rather than  $(2^2) + (2^3) = 12$ . This means that arithmetic order of operations is not followed in multi-operation terms. This is a limitation and is discussed in the future work section.

The final limitation to be mentioned is the lack of responsive design in the interface. This means that elements on screen may not display correctly on some devices. This could be easily remedied in future but has not been done due to time restrictions, this has again been discussed in the future work section.

## 7.2 Improving Student Understanding

The user tests conducted and summarised in section 6.2 suggest that students with a computer science background learning lambda calculus for the first time would likely find the tool useful in supporting their learning. Possible improvements such as including tooltips as suggested in user feedback are valid and would be implemented given more time.

The user feedback also suggested that the simplicity of the interface is a positive, and care must be taken to prevent the interface from becoming too complicated. With a goal of supporting learning by providing an interface which accurately evaluates lambda terms, it can be said that these goals have been met.

## 7.3 Future Work

Due to the time limited nature of this project, not all features or corrections have been implemented. Given more time, the following work would be done:

1. Fix the order of arithmetic operations issue discussed in the Limitations section of this report, this was simply left out due to time restrictions with the project
2. Add in multi-input abstraction terms to allow users to enter  $\lambda xy.M$  as appears in the lecture slides [8] as opposed to  $\lambda x.\lambda y.M$ , which would allow students to directly copy more lambda terms from the lecture slides to see how they are evaluated.
3. Improve the web interface by implementing responsive design

4. Improve the type inference to allow input types of multi-operation functions to be inferred.
5. Include tooltips linked to various aspects of the web interface to give users more information about what particular phrases mean or what elements of the web interface do.
6. Do further user testing for the web interface to get a wider understanding of any style/usability issues that exist from a wider demographic of participants. This could include those who have previously taken the Theory of Computation course to identify any areas which would most benefit from further support, those learning lambda calculus using alternative methods, or those who have less of a background in computer science than the participants used in the study.
7. Allow users to enter lambda terms in the De Bruijn notation, for those unfamiliar with the Barendregt convention or those wanting to expand their knowledge of Lambda Calculus.
8. Add a feature which would evaluate the term step-by-step, so evaluation could be paused. This would allow users to see exactly how the term is being evaluated, and the state of evaluation at each stage.

## Chapter 8: Conclusion

This report has documented the work done in creating a tool to evaluate lambda terms designed to support those studying the Theory of Computation at Glasgow University. The report has given an introduction to key terms, an analysis to determine specific objectives, and has detailed the design and implementation of the tool. It has also covered the testing and evaluation of the final interface to determine how well key goals have been met.

The outcome of the project was extremely successful, resulting in a simple and convenient tool which accurately evaluates lambda calculus terms in an easy-to-use web interface. Design decisions were made based on lecture material, but the tool was not limited so much that it could only be used by those on the Theory of Computation course. The final tool is broad enough to be accessible to those learning lambda calculus through alternate methods, but specific enough to support those on the course achieve understanding of key learning outcomes.

Creating this tool was very much a challenge, having no prior knowledge of lambda calculus before the project was started resulted in an extremely steep learning curve and a number of setbacks due to misunderstandings made early on in the project. As well as this ANTLR had to be learned from scratch, resulting in a portion of time being allocated to becoming familiar with the tool in order to utilise it in the best way possible.

Due to the finite nature of the project and the amount of learning that had to be done as implementation was underway, there are features which would benefit the tool which have not been included due to time restrictions. These include supporting multiple notations such as De Bruijn notation and multi-input terms which have been discussed previously.

The tool could also be expanded in ways which would further benefit student learning, for example adding tooltips to give users more guidance on the theoretical aspects of lambda calculus and allowing students to evaluate terms step-by-step so they could see how the term is processed as evaluation is happening.

As well as additional features, one key goal would be to perform additional user testing across a wider demographic of participants, including those who previously have taken the Theory of Computation course, those more familiar with lambda calculus and those at other universities or learning lambda calculus through other means. This would help improve the tool for a wider range of potential users, and allow additional feature ideas to come from people who would actually be using it.

However, despite the narrow demographic of user feedback and the additional features which could potentially be implemented, the project was a huge success. From the 12 test participants, 100% said they found the tool useful and would use it if they were on the Theory of Computation course, reporting a 39% improvement in understanding having used the tool for a short period of time. Testing has also shown that the tool does evaluate lambda terms accurately, by following the rules of lambda calculus as closely as possible and making the best use of the extremely powerful parsing tool that is ANTLR. This has resulted in a system which will be accessible to students on the Theory of Computation course next year, being hosted on university servers in the near future.

Creating the tool was definitely a challenge, due to having no knowledge of lambda calculus or ANTLR before starting the project. But the tool created has met all the goals which were set out at the start of the project, and includes additional features going beyond the limited specification to further improve student understanding. It is hoped that future students will find this tool useful in their university studies, and that it will help to improve their understanding of lambda calculus and functional programming.

# Bibliography

- [1] Umut Acar and Amal Ahmed. *Lecture 3: Lambda Calculus (Syntax, Substitution, Beta Reduction)*. Jan. 17, 2018. URL: <https://ttic.uchicago.edu/~pl/classes/CMSC336-Winter08/lectures/lec3.pdf> (visited on 08/01/2019).
- [2] Henk Barendregt and Erik Barendsen. *Introduction to Lambda Calculus*. Mar. 2000. URL: <http://www.cse.chalmers.se/research/group/logic/TypesSS05/Extra/geuvers.pdf> (visited on 07/27/2019).
- [3] Marc Bezem and Jan F. Groote. *Typed Lambda Calculi and Applications*. Springer-Verlag, June 13, 2008, p. 356.
- [4] Steven Bird, Ewan Klein, and Edward Loper. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. O'Reilly, 2009, pp. 291–322.
- [5] Rosalie Chan. *The 10 most popular programming languages, according to the 'Facebook for programmers'*. Aug. 10, 2019. URL: <https://www.businessinsider.com/the-10-most-popular-programming-languages-according-to-github-2018-10?r=US&IR=T> (visited on 01/22/2019).
- [6] Alonzo Church and J. B. Rosser. “Some Properties of Conversion”. In: *Transactions of the American Mathematical Society* 38 (3 1936), pp. 472–482.
- [7] Agile Business Consortium. *The DSDM Agile Project Framework*. Jan. 2014.
- [8] Ornela Dardha and Simon Gay. *The Theory of Computation (H), course code COMP-SCI4072 Lecture Slides*. 2019. URL: <https://www.gla.ac.uk/coursecatalogue/course/?code=COMPSCI4072>.
- [9] Stacy Everett. *Grammar*. July 11, 2019. URL: <https://github.com/antlr/grammars-v4/tree/master/arithmetic> (visited on 07/28/2019).
- [10] Python Software Foundation. *Built-in Exceptions*. Aug. 8, 2019. URL: <https://docs.python.org/3/library/exceptions.html> (visited on 08/08/2019).
- [11] University of Glasgow. *Theory of Computation (H) Course Specification*. 2019. URL: <https://www.gla.ac.uk/coursecatalogue/document?type=sp&courseCode=COMPSCI4072> (visited on 08/02/2019).
- [12] Chris Hankin. *An Introduction to Lambda Calculi for Computer Scientists*. 2nd ed. King's College Publication, 2004.
- [13] Mehedia Hasan. *The 20 Most Popular Programming Languages To Learn For Your Open Source Project*. Aug. 10, 2019. URL: <https://www.ubuntupit.com/top-20-most-popular-programming-languages-to-learn-for-your-open-source-project/> (visited on 2019).
- [14] Susan B. Horowitz. *Lambda Calculus (Part 1)*. 2013. URL: <http://pages.cs.wisc.edu/~horowitz/CS704-NOTES/1.LAMBDA-CALCULUS.html#normal> (visited on 08/01/2019).
- [15] Paul Hudak. *A Brief and Informal Introduction to the Lambda Calculus*. 2008. URL: <http://www.cs.yale.edu/homes/hudak/CS201S08/lambda.pdf> (visited on 08/20/2019).

- [16] Fairouz Kamareddine. *Reviewing the Classical and the de Bruijn Notation for  $\lambda$ -calculus and Pure Type Systems*. Oct. 4, 2000. URL: <https://pdfs.semanticscholar.org/2b65/df6c8c0cc14db2afcacbca0f0423588ec6f0.pdf>? (visited on 08/02/2019).
- [17] Will Kenton. *T-Test*. Jan. 25, 2019. URL: <https://www.investopedia.com/terms/t/t-test.asp> (visited on 08/22/2019).
- [18] Wayne W. LaMorte. *Hypothesis Testing: Upper-, Lower, and Two Tailed Tests*. Nov. 6, 2017. URL: [http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704\\_HypothesisTest-Means-Proportions/BS704\\_HypothesisTest-Means-Proportions3.html](http://sphweb.bumc.bu.edu/otlt/MPH-Modules/BS/BS704_HypothesisTest-Means-Proportions/BS704_HypothesisTest-Means-Proportions3.html) (visited on 08/22/2019).
- [19] Timothy C. Lethbridge and Robert Laganière. *Object-Oriented Software Engineering*. McGraw Hill Book Company, Dec. 1, 2004, p. 323.
- [20] Ralph Loader. *Notes on Simply Typed Lambda Calculus*. Feb. 1998. URL: <http://www.lfcs.inf.ed.ac.uk/reports/98/ECS-LFCS-98-381/ECS-LFCS-98-381.pdf> (visited on 07/28/2019).
- [21] Matt Makai. *Flask*. Aug. 8, 2019. URL: <https://www.fullstackpython.com/flask.html> (visited on 2019).
- [22] Greg Michaelson. *An Introduction to Functional Programming through Lambda Calculus*. Dover Publications Inc., Apr. 21, 2011.
- [23] Terrance Parr. *The Definitive Antlr4 Reference*. 1st ed. The Pragmatic Programmers, 2012.
- [24] Alison DeNisco Payome. *Forget the most popular programming languages, here's what developers actually use*. Aug. 10, 2019. URL: <https://www.techrepublic.com/article/forget-the-most-popular-programming-languages-heres-what-developers-actually-use/> (visited on 10/22/2018).
- [25] Benjamin C. Pierce. *Types and Programming Languages*. The MIT Press, 2002.
- [26] Raúl Rojas. *A Tutorial Introduction to the Lambda Calculus*. 1998. URL: <https://www.inf.fu-berlin.de/lehre/WS03/alpi/lambda.pdf> (visited on 07/28/2019).
- [27] Peter Selinger. *Lecture Notes on the Lambda Calculus*. 2013. URL: <https://www.irif.fr/~mellies/mpri/mpri-ens/biblio/Selinger-Lambda-Calculus-Notes.pdf> (visited on 07/27/2019).
- [28] Ken Slonneger and Barry Kurtz. *Chapter 5: The Lambda Calculus*. Pearson, 1995, pp. 139–166.
- [29] Kevin Sookocheff. *Eta Reduction*. Sept. 27, 2018. URL: <https://sookocheff.com/post/fp/eta-conversion/> (visited on 08/02/2019).
- [30] Kevin Sookocheff. *Normal, Applicative and Lazy Evaluation*. Sept. 18, 2018. URL: <https://sookocheff.com/post/fp/evaluating-lambda-expressions/> (visited on 08/01/2019).
- [31] Saumitra Srivastav. *Antlr4 - Visitor vs Listener Pattern*. Oct. 19, 2017. URL: <https://saumitra.me/blog/antlr4-visitor-vs-listener-pattern/> (visited on 08/06/2019).
- [32] Paul Tarau. *Church Encoding*. Feb. 17, 2017. URL: <http://www.cse.unt.edu/~tarau/teaching/PL/docs/Church%20encoding.pdf> (visited on 08/02/2019).
- [33] Gabriele Tomassetti. *The ANTLR Mega Tutorial*. Mar. 8, 2007. URL: <https://tomassetti.me/antlr-mega-tutorial/> (visited on 08/06/2019).

- [34] Sebastian Wiesner. *Type Inference - Main seminar, summer term 2011*. Aug. 14, 2019.  
URL: <http://www2.in.tum.de/hp/file?fid=880> (visited on 06/22/2011).