



# 60023

**Type Systems for  
Programming Languages  
Imperial College London**

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Lambda Calculus</b>	<b>3</b>
2.1	Introduction to Lambda Calculus . . . . .	3
2.2	Reduction Strategies . . . . .	5
2.2.1	Head Reduction . . . . .	5
2.2.2	Call By Name / Lazy . . . . .	5
2.2.3	Call By Value . . . . .	5
2.2.4	Normal Order . . . . .	5
2.2.5	Applicative Order . . . . .	5
2.2.6	Computability . . . . .	6
2.3	Normal Forms . . . . .	6
2.4	Approximation Semantics . . . . .	8
2.4.1	Properties of Approximants . . . . .	9
<b>3</b>	<b>Curry Type Assignment</b>	<b>10</b>
3.0.1	Curry Type Assignment . . . . .	10
3.0.2	Important Lemmas For Type Assignment . . . . .	11
3.1	Principle Type Property . . . . .	11
3.1.1	Unification . . . . .	12
3.1.2	Curry Principle Pair . . . . .	12
<b>4</b>	<b>Polymorphism</b>	<b>13</b>
4.1	Language $\Lambda^N$ . . . . .	13
4.2	Type Assignment for $\Lambda^N$ . . . . .	14
4.3	Principal Types for $\Lambda^N$ . . . . .	14
<b>5</b>	<b>Recursion</b>	<b>16</b>
5.1	Language $\Lambda^{NR}$ . . . . .	16
5.2	Type Assignment for $\Lambda^{NR}$ . . . . .	16
5.3	Principle Types for $\Lambda^{NR}$ . . . . .	17
<b>6</b>	<b>Milner's ML</b>	<b>18</b>
6.1	The ML Type Assignment System . . . . .	18
6.1.1	Term Substitution . . . . .	18
6.1.2	Reduction . . . . .	19
6.1.3	Type Assignment . . . . .	19
6.1.4	Lemmas for Type Assignment . . . . .	20
6.2	Milner's $\mathcal{W}$ . . . . .	22
6.2.1	Basic Cases . . . . .	22
6.2.2	Let Construct . . . . .	23
6.2.3	Fix Construct . . . . .	23
6.2.4	Application . . . . .	23
6.3	Polymorphic Recursion . . . . .	23
6.3.1	Mycroft-Style Assignment for $\Lambda^{NR}$ . . . . .	24
6.3.2	Milner's and Mycroft's System's Differences . . . . .	24
<b>7</b>	<b>Credit</b>	<b>26</b>

# Chapter 1

## Introduction

# Chapter 2

## Lambda Calculus

### 2.1 Introduction to Lambda Calculus

$\lambda$ -Terms	Definition 2.1.1
<p>Given the set of term-variables <math>\mathcal{V} = \{x, y, z, \dots\}</math>, a <math>\lambda</math>-term is defined by the grammar:</p> $M, N ::= \underset{\text{variable}}{x} \quad   \quad \underset{\text{abstraction}}{(\lambda x.M)} \quad   \quad \underset{\text{application}}{(M \ N)}$ <p>We can also describe this using an <i>inference system</i>:</p> $\frac{}{x \in \Lambda} (x \in \mathcal{V}) \quad \frac{M \in \Lambda}{(\lambda x.M) \in \Lambda} ((x \in \mathcal{V})) \quad \frac{M \in \Lambda \quad N \in \Lambda}{(M \ N) \in \Lambda}$ <ul style="list-style-type: none"> <li>• In a lambda term <math>M \cdot N</math>, <math>M</math> is in the <i>function position</i> and <math>N</math> is an <i>argument</i></li> <li>• The leftmost, outer brackets can be ommitted (<math>M \ N \ (P \ Q) = ((M \ N) \ (P \ Q))</math>)</li> <li>• Abstractions can be abbreviated <math>\lambda xyz.M = (\lambda x.(\lambda y.(\lambda z.M)))</math></li> <li>• Computation is expressed through term substitution.</li> </ul>	

Free Variables	Definition 2.1.2	Bound Variables	Definition 2.1.3
$\begin{aligned} fv(x) &= \{x\} \\ fv(\lambda y.M) &= fb(M) \setminus \{y\} \\ fb(M \ N) &= fv(M) \cup fv(N) \end{aligned}$ <p>A <math>\lambda</math>-term <math>M</math> is closed if <math>fv(M) = \emptyset</math>.</p>		$\begin{aligned} bv(x) &= \emptyset \\ bv(\lambda y.M) &= bv(M) \cup y \\ bv(M \ N) &= bv(M) \cup bv(N) \end{aligned}$ <p>A term with no free variables is <i>closed</i>.</p>	

We can define term substitution inductively as:  
Where  $P[N/x]$  means replace  $x$  by  $N$  in  $\lambda$ -term  $P$ .

This definition can result in variable capture, for example:

$$\begin{aligned} x[N/x] &= N \\ y[N/x] &= y \\ (P \ Q)[N/x] &= P[N/x] \ Q[N/x] \\ (\lambda y.M)[N/x] &= \lambda y.(M[N/x]) \text{ where } y \neq x \\ (\lambda x.M)[N/x] &= \lambda x.M \end{aligned}$$

$$(\lambda x.y \ x)[y/x] = \lambda x.x \ x$$

Here the free  $y$  was substituted for another free variable  $x$ , however has been captured by the bound  $x$  in the abstraction.

Barendregt's convention	Definition 2.1.4
<p>Given some <math>(\lambda x.M)N</math> we can assume:</p> $\begin{aligned} x \notin fv(N) & \quad x \text{ is not free in } N \\ \forall y \in bv(M). [y \notin fv(N)] & \quad \text{All bound variables in } M \text{ are not free in } N \end{aligned}$ <p>We can always rename the bound variables of a term, this is a fundamental feature to the degree that <math>\alpha</math>-conversion rarely plays a role and terms are considered modulo <math>\alpha</math>-conversion.</p>	

Equivalence Relation	Definition 2.1.5
A binary relation that is reflexive, symmetric and transitive.	

$\alpha$ -Conversion	Definition 2.1.6	$\alpha$ -Equivalence	Definition 2.1.7
$(\lambda x.M)N \rightarrow_\alpha (\lambda z.M[z/x])N$ where $z$ is a new Renaming bound variables within a term.		$N \rightarrow_\alpha M \wedge M \rightarrow_\alpha N \Leftrightarrow M =_\alpha N$ Terms that can be made equal by $\alpha$ -conversion are $\alpha$ -Equivalent	

$\beta$ -Conversion	Definition 2.1.8
$\begin{array}{ccc} (\lambda x.M)N & \rightarrow_\beta & M[N/x] \\ \text{Reducible Expression/Redex} & & \text{Contractum/Reduct} \end{array}$ <p>The <i>one-step</i> reduction <math>\rightarrow_\beta</math> can be defined with contextual closure rules:</p> $M \rightarrow_\beta N \Rightarrow \begin{cases} \lambda x.M & \rightarrow_\beta \lambda x.N \\ P M & \rightarrow_\beta P N \\ M P & \rightarrow_\beta N P \end{cases}$ <p><math>\rightarrow^*_\beta</math> or <math>\rightarrow_\beta</math> is the transitive closure of <math>\rightarrow_\beta</math>. We can also define this using an inference system:</p> $\begin{aligned} (\beta) : \frac{}{(\lambda x.M)N \rightarrow_\beta M[N/x]} \quad (\text{Appl-L}) : \frac{M \rightarrow_\beta N}{M P \rightarrow_\beta N P} \quad (\text{Appl-R}) : \frac{M \rightarrow_\beta N}{P M \rightarrow_\beta P N} \\ (\text{Abstr}) : \frac{M \rightarrow_\beta N}{\lambda x.M \rightarrow_\beta \lambda x.N} \\ (\text{Inherit}_r) : \frac{M \rightarrow_\beta N}{M \rightarrow^*_\beta N} \quad (\text{Refl}) : \frac{}{M \rightarrow^*_\beta M} \quad (\text{Trans}_r) : \frac{M \rightarrow^*_\beta N \quad N \rightarrow^*_\beta P}{M \rightarrow^*_\beta P} \\ (\text{Inherit}_l) : \frac{M \rightarrow^*_\beta N}{M =_\beta N} \quad (\text{Symm}) : \frac{M =_\beta N}{N =_\beta M} \quad (\text{Trans}_{eq}) : \frac{M =_\beta N \quad N =_\beta P}{M =_\beta P} \end{aligned}$ <p><math>\beta</math>-reduction is confluent/satisfies the Church-Rosser property:</p> $\forall N, M, P. [M \rightarrow^*_\beta N \wedge M \rightarrow^*_\beta P \Rightarrow \exists Q. [N \rightarrow^*_\beta Q \wedge P \rightarrow^*_\beta Q]]$	

$\beta$ -conversion does not conform to *Barendregt's convention*, for example:

$$\begin{aligned} (\lambda xy.xy)(\lambda xy.xy) &\rightarrow (\lambda xy.xy)[(\lambda xy.xy)/x] = \lambda y.(\lambda xy.xy)y \\ &\rightarrow \lambda y.(\lambda xy.xy)[y/x] = \lambda y.(\lambda y.yy) \end{aligned}$$

We can avoid this by alpha converting the term to  $\lambda y.(\lambda xz.xz)y$  before  $\beta$ -conversion.

$\eta$ -Reduction	Definition 2.1.9
<p>Given <math>x \notin fv(M)</math> then <math>\lambda x.M x \rightarrow_\eta M</math></p> <p><math>\eta</math>-reduction can be used for eta equivalence. If <math>f x = g x</math> then we can eta reduce both to <math>f = g</math>.</p> <ul style="list-style-type: none"> <li>• Eta reduction is a common lint provided by <code>hlint</code> for <code>haskell</code>.</li> </ul>	

## 2.2 Reduction Strategies

Evaluation Context	Definition 2.2.1
<p>A term with a single hole <math>\square</math> :</p> $C ::= \square \mid C M \mid M C \mid \lambda x.C$ <p><math>C[M]</math> is the term obtained from context <math>C</math> by replacing the <i>hole</i> <math>\square</math> with <math>M</math>.</p> <ul style="list-style-type: none"> <li>This allows any variables to be captured.</li> </ul> <p>The one step <math>\beta</math>-reduction rule can be defined for any evaluation context as:</p> $C_N[(\lambda x.M)N] \rightarrow C_N[M[N/x]]$	

### 2.2.1 Head Reduction

$$\frac{}{(\lambda x.M)N \rightarrow_H M[N/x]} \quad \frac{M \rightarrow_H N}{\lambda x.M \rightarrow_H \lambda x.N} \quad \frac{M \rightarrow_H N}{M P \rightarrow_H N P}$$

Reduce the leftmost term, if this is an abstraction, reduce the inside of the abstraction.

### 2.2.2 Call By Name / Lazy

$$\frac{}{(\lambda x.M)N \rightarrow_N M[N/x]} \quad \frac{M \rightarrow_N N}{M P \rightarrow_N N P}$$

Reduce the leftmost term. Do not reduce unless a term is applied (lazy evaluation).

We can also express reduction strategy with an evaluation context:

$$C_N ::= \square \mid C_N M \quad \text{where } \rightarrow_\beta^N \text{ is defined as } C_N[(\lambda x.M)N] \rightarrow C_N[M[N/x]]$$

Note that there is only ever one redex to contract.

### 2.2.3 Call By Value

Given  $V$  denotes abstractions and variables (values):

$$\frac{}{(\lambda x.M)V \rightarrow_V M[V/x]} \quad \frac{M \rightarrow_V N}{M P \rightarrow_V N P} \quad \frac{M \rightarrow_V N}{V M \rightarrow_V V N}$$

We can apply values, the leftmost term that is not a value is reduced first.

We can also express reduction strategy with an evaluation context:

$$C_V ::= \square \mid C_V M \mid V C_V \quad \text{where } \rightarrow_\beta^V \text{ is defined as } C_V[(\lambda x.M)V] \rightarrow C_V[M[V/x]]$$

Note that there is only ever one redex to contract.

### 2.2.4 Normal Order

$$\frac{}{(\lambda x.M)N \rightarrow_N M[N/x]} \quad \frac{M \rightarrow_N N}{M P \rightarrow_N N P} \quad \frac{M \rightarrow_N N}{P M \rightarrow_N P N} (P \text{ contains no redexes}) \quad \frac{M \rightarrow_N N}{\lambda x.M \rightarrow_N \lambda x.N}$$

Reduce the leftmost term until it contains no redexes (then continue to other terms), can reduce the inside of an abstraction.

### 2.2.5 Applicative Order

$$\frac{}{(\lambda x.M)N \rightarrow_A M[N/x]} (M, N \text{ contain no redexes}) \quad \frac{M \rightarrow_A N}{M P \rightarrow_a N P} \\ \frac{M \rightarrow_A N}{P M \rightarrow_A P N} (P \text{ contains no redex}) \quad \frac{M \rightarrow_A N}{\lambda x.M \rightarrow_A \lambda x.N}$$

## 2.2.6 Computability

SKI Combinator Calculus	Definition 2.2.2
$\mathcal{S} = \lambda xyz.xz(yz) \quad \mathcal{K} = \lambda xy.x \quad \mathcal{I} = \lambda x.x$ <p>Any operation in lambda calculus can be encoded (by <i>abstraction elimination</i>) into the SKI calculus as a binary tree with leaves of symbols <math>\mathcal{S}</math>, <math>\mathcal{K}</math> &amp; <math>\mathcal{I}</math>.</p>	

It is possible to encode all Turing Machines within *lambda*-calculus and vice versa. This makes  $\lambda$ -calculus (along with Turing Machines) a model for what is computable.

Church-Turing thesis	Extra Fun! 2.2.1
<p>The Church-Turing thesis equivocates the computational power of Turing machines and the lambda calculus. (Wikipedia)</p>	

It is possible to write terms that do not terminate under  $\beta$ -reduction:

$$(\lambda x.xx) (\lambda x.xx) \rightarrow_{\beta} (xx)[(\lambda x.xx)/x] = (\lambda x.xx) (\lambda x.xx)$$

We can also apply functions continuously.

$$\begin{aligned} \lambda f.(\lambda x.f(x x))(\lambda x.f(x x)) &\rightarrow_{\beta} \lambda f.f(x x)[(\lambda x.f(x x))/x] &&= \lambda f.f((\lambda x.f(x x))(\lambda x.f(x x))) \\ &\rightarrow_{\beta} \lambda f.f(f((\lambda x.f(x x))(\lambda x.f(x x)))) \\ &\rightarrow_{\beta} \lambda f.f(f(f((\lambda x.f(x x))(\lambda x.f(x x)))))) \\ &\vdots \\ &\rightarrow_{\beta} \lambda f.f(f(f(f(f(\dots))))) \end{aligned}$$

This term is a *fixed point constructor*.

Fixed-Point Theorem	Definition 2.2.3
$\forall M.\exists N.[M N =_{\beta} N]$ <p>Take <math>N = Y M</math> where <math>Y = \lambda f.(\lambda x.f(x x))(\lambda x.f(x x))</math>:</p> $\begin{aligned} Y M &\triangleq \lambda f.(\lambda x.f(x x))(\lambda x.f(x x)) M \\ &\rightarrow_{\beta} (\lambda x.M(x x))(\lambda x, M(x x)) \\ &\rightarrow_{\beta} (\lambda x.M(x x))(\lambda x, M(x x)) \\ M(Y M) &\triangleq M(\lambda f.(\lambda x.f(x x))(\lambda x.f(x x)) M) \\ &\rightarrow_{\beta} M((\lambda x.M(x x))(\lambda x, M(x x))) \end{aligned}$ <p>Hence <math>M(Y M) =_{\beta} Y M</math> meaning that <math>Y</math> is the fixed point constructor of <math>M</math></p>	

## 2.3 Normal Forms

Normal Form	Definition 2.3.1
<p>A <math>\lambda</math>-term is in normal form if it does not contain a <i>redex</i>.</p> $N ::= x \mid \lambda x.N \mid xN_1 \dots N_n \text{ where } (n \geq 0)$ <p>No <math>\beta</math> or <math>\eta</math> reductions are possible</p>	

Head Normal Form		Definition 2.3.2
A $\lambda$ -term is in head normal form if it is an abstraction with a body that is not <i>reducible</i> .		
$H ::= x \mid \lambda x.H \mid xM_1 \dots M_n \text{ where } n \geq 1 \wedge M_i \in \Lambda$		
This will mean the term is of the form $x$ or $\lambda x_1 \dots x_n.yM_1 \dots M_m$		
<ul style="list-style-type: none"> <li>• <math>y</math> is the <i>head-variable</i></li> <li>• If a term has a head-normal form, then head-reduction on the term terminates.</li> </ul>		
Head Normalisable	Definition 2.3.3	Strongly Normalisable
A term $M$ is head normalisable if it has a head-normal form.		Definition 2.3.4
$M \rightarrow_\beta^* N$ where $N$ is in head normal form		A term $M$ is strongly normalisable if all reduction sequences starting from $M$ are finite.
Meaningless		Definition 2.3.5
A term without a head-normal form is meaningless as it can never interact with any context (can never apply it to some argument).		
Normal Forms		Example Question 2.3.1
Determine the normality of the following terms:		
<ol style="list-style-type: none"> <li>1. <math>\lambda f.(\lambda x.f(x\ x))\ (\lambda x.f(x\ x))</math></li> <li>2. <math>(\lambda x.x\ x)\ (\lambda x.x\ x)</math></li> <li>3. <math>\mathcal{S}\ \mathcal{K}</math></li> <li>4. <math>(\lambda a b.b)\ ((\lambda x.x\ x)\ (\lambda x.x\ x))</math></li> </ol>		
<ol style="list-style-type: none"> <li>1. Not in either head normal form or normal form (contains a redex). <div style="text-align: center;"> <math display="block">\lambda f.(\lambda x.f(x\ x))(\lambda x.f(x\ x))</math> <math display="block">\rightarrow_\beta \lambda f.f((\lambda x.f(x\ x))\ (\lambda x.f(x\ x)))</math> </div> <p>However the <math>\beta</math>-reduction is in head normal form (head-variable is <math>f</math>).</p> </li> <li>2. It is a redex, so its not in a normal form. It does not have a normal form as it reduces to itself, so all reducts contain a redex. It has no head-normal form.</li> <li>3. Hence the original <math>\lambda</math>-term is not normal form, but it can be normalised. <div style="margin-left: 40px;"> <math display="block">\begin{aligned} &amp; \mathcal{S}\ \mathcal{K} &amp;&amp; \text{Must expand } \mathcal{S} \text{ and } \mathcal{K} \\ = &amp; (\lambda x y z.xz(yz))\ (\lambda x y.x) &amp;&amp; \text{Is a redex} \\ \rightarrow_\beta &amp; (\lambda x y z.xz(yz))\ (\lambda x y.x) &amp;&amp; \text{We rename } y \text{ as per barendregt's convention} \\ =_\alpha &amp; (\lambda x y z.xz(yz))\ (\lambda x a.x) \\ \rightarrow_\beta &amp; (\lambda y z.(\lambda x a.x)z(yz)) \\ \rightarrow_\beta &amp; (\lambda y z.(\lambda a.z)(yz)) \\ \rightarrow_\beta &amp; (\lambda y z.z) \end{aligned}</math> </div> <p>As all possible redexes are contracted it is <i>strongly normalisable</i>.</p> </li> <li>4. Contracting the outermost redex results in normal form ter <math>\lambda b.b</math>. However contracting the inner term yields itself. Hence it is normalisable, but not <i>strongly normalisable</i>.</li> </ol>		



## 2.4 Approximation Semantics

There are many methods of describing the semantics of the  $\lambda$ -calculus.

- Reduction rules with *operational semantics*
- set theory with *denotational semantics*

The approach studied in this module defines semantics in a denotational style, but using a reduction system for its definition.

We introduce an extension to the  $\lambda$ -calculus syntax by adding the constant  $\perp$ ,

- $\perp$  means unknown/meaningless/no information
- used to mask sub-terms (typically containing redexes) to allow us to focus on the the *stable* parts of the term that do not change under reduction.

The set of  $\Lambda\perp$ -terms is defined as:

$$M, N ::= z \mid \perp \mid \lambda x.M \mid M N$$

$\beta$ -reduction is extended to  $\rightarrow_\perp$  to include:

$$\lambda x.\perp \rightarrow_\perp \perp \quad \text{and} \quad \perp M \rightarrow_\perp \perp$$

The set of normal forms of  $\Lambda\perp$  with respect to  $\rightarrow_\perp$  is the set  $\mathcal{A}$ :

$$A ::= \perp \mid \lambda x.A \ (A \neq \perp) \mid xA_1 \dots A_n$$

Note that  $\lambda x.\perp$  is considered a redex.

### Approximant

### Definition 2.4.1

An approximant is a redex-free  $\Lambda\perp$ -normal forms that can contain  $\perp$  and are used to represent finite parts of potentially infinitely large  $\lambda$ -terms in head-normal form.

The partial order  $\sqsubseteq \subseteq (\Lambda\perp)^2$  is defined as the smallest pre-order (reflexive and transitive) such that:

$$\begin{array}{lll} \perp \sqsubseteq M & M \sqsubseteq M' & \Rightarrow \lambda x.M \sqsubseteq \lambda x.M' \\ x \sqsubseteq x & M_1 \sqsubseteq M'_1 \wedge M_2 \sqsubseteq M'_2 & \Rightarrow M_1 M_2 \sqsubseteq M'_1 M'_2 \end{array}$$

- For  $A \in \mathcal{A}, M \in \Lambda$ , if  $A \sqsubseteq M$  then  $A$  is the *direct approximant* of  $M$
- The set of *approximants* of  $M$ ,  $\mathcal{A}(M)$  is defined as:

$$\mathcal{A}(M) \triangleq \{A \in \mathcal{A} \mid \exists M' \in \Lambda. [M \rightarrow_\beta^* M' \wedge A \sqsubseteq M']\}$$

- If  $A$  is a *direct approximant* of  $M$ , then  $A$  and  $M$  have the same structure, but some parts  $A$  contains  $\perp$  ( $\perp$  masking part of  $M$ ).
- Redexes in  $M$  are masked by  $\perp$  in  $A$  ( $\perp$  masks the redex, or a larger location that contains the redex).

### Direct Approximants

### Example Question 2.4.1

Show the direct approximants for each reduction step of:

1.  $\mathcal{S} \mathcal{K}$
2.  $\mathcal{S} a \mathcal{K}$

1.

$$\mathcal{S} \mathcal{K} = (\lambda xyz.xz(yz)) (\lambda ab.a) \xrightarrow{\beta} \lambda yz.(\lambda ab.a)z(yz) \xrightarrow{\beta} \lambda yz.(\lambda b.z)(yz) \xrightarrow{\beta} \lambda yz.z$$

$$\{\perp\} \quad \{\perp\} \quad \{\perp\} \quad \{\perp, \lambda yz.z\}$$

2.

$$\begin{array}{lll} \mathcal{S} a \mathcal{K} = & (\lambda xyz.xz(yz)) a (\lambda cd.c) & \{\perp\} \\ \xrightarrow{\beta} & (\lambda yz.az(yz)) (\lambda cd.c) & \{\perp\} \\ \xrightarrow{\beta} & (\lambda z.az((\lambda cd.c)z)) & \{\perp, \lambda z.a\perp\perp, \lambda z.az\perp\} \\ \xrightarrow{\beta} & (\lambda z.az(\lambda d.z)) & \{\perp, \lambda z.a\perp\perp, \lambda z.az\perp, \lambda a\perp(\lambda d.z), \lambda az(\lambda d.z)\} \end{array}$$

Some basic approximants are:

$$\begin{aligned}
\mathcal{A}(\lambda x.x) &= \{\perp, \lambda x.x\} \\
\mathcal{A}(\lambda x.x x) &= \{\perp, \lambda x.x \perp, \lambda x.x x\} \\
\mathcal{A}(\lambda x.x((\lambda y.yy)(\lambda y.yy))) &= \{\perp, \lambda x.x \perp\} \\
\mathcal{A}(\mathcal{S}) = \mathcal{A}(\lambda xyz.xz(yz)) &= \{\perp, \lambda xyz.x \perp \perp, \lambda xyz.x \perp(y \perp), \lambda xyz.x \perp(yz), \lambda xyz.xz \perp, \lambda xyz.xz(y \perp), \lambda xyz.xz(yz)\} \\
\mathcal{A}(\lambda f.(\lambda x.f(x x)) (\lambda x.f(x x))) &= \{\perp, \lambda f.f(\perp), \lambda f.f(f(\perp)), \lambda f.f(f(f(\perp))), \dots\}
\end{aligned}$$

### 2.4.1 Properties of Approximants

$$(A \in \mathcal{A}(xM_1 \dots M_n) \wedge A \neq \perp \wedge A' \in \mathcal{A}(N)) \Rightarrow AA' \in \mathcal{A}(xM_1 \dots M_n N)$$

Given  $A$  is in the approximants of some variable  $x$  are lambda terms  $M_1 \dots M_n$ , and  $A'$  in the approximants of  $N$ , then  $AA'$  is in the approximants of  $A A'$  (Applying  $A$  to  $A'$ ).

$$(A \in \mathcal{A}(Mz) \wedge z \notin fv(M)) \Rightarrow \left( \begin{array}{l} A = \perp \\ \vee \quad A \equiv A'z \text{ where } z \notin fv(A') \wedge A' \in \mathcal{A}(M) \\ \vee \quad \lambda x.A \in \mathcal{A}(M) \end{array} \right)$$

If  $A$  is an approximant of  $Mz$ , and  $z$  is not free in  $M$ , then either:

- $A$  is  $\perp$
- $A$  is some  $A'z$ , hence by  $\eta$ -reduction, we can see  $A' \in \mathcal{A}(M)$  (the  $z$  part can be disregarded, and just look at approximants of  $M$ ).

$$A \sqsubseteq M \wedge M \rightarrow_{\beta}^* N \Rightarrow A \sqsubseteq N$$

If  $A$  is ordered before  $M$ , and  $M$   $\beta$ -reduces to  $N$ , then  $A$  is also before  $N$ .

$$A \in \mathcal{A}(M) \wedge M \rightarrow_{\beta}^* N \Rightarrow A \in \mathcal{A}(N) \quad A \in \mathcal{A}(N) \wedge M \rightarrow_{\beta}^* N \Rightarrow A \in \mathcal{A}(M)$$

**UNFINISHED!!!**

## Chapter 3

# Curry Type Assignment

Type assignment follows the syntactic structure of terms. For example  $\lambda x.M$  will be of the form  $A \rightarrow B$  where the input  $x$  is of type  $A$ , and  $M$  is of type  $B$ .

$\mathcal{T}_C$  is the set of *types*.

- This is ranged over by  $A, B \dots$  and defined over the set of *type variables*  $\Phi$ .
- The set of *type variables*  $\Phi$  is ranged over by  $\varphi$

$$A, B ::= \varphi \mid (A \rightarrow B)$$

A type can be either some type variable (some type e.g Int), or a function converting one type to another.

Statement	Definition 3.0.1
An expression of the form $M : A$ where $M \in \Lambda$ and $A \in \mathcal{T}_C$ .	
<ul style="list-style-type: none"> <li>• <math>M</math> is the <i>subject</i></li> <li>• <math>A</math> is the <i>predicate</i></li> </ul>	

Context	Definition 3.0.2
A context $\Gamma$ is a set of statements with distinct variables as subjects.	
<ul style="list-style-type: none"> <li>• <math>\Gamma, x : A</math> is shorthand for <math>\Gamma \cup \{x : A\}</math> where <math>x</math> does not occur as a subject in <math>\Gamma</math> (variables must be distinct).</li> <li>• <math>x : A</math> is shorthand for <math>\emptyset, x : A</math>.</li> <li>• <math>x \in \Gamma</math> is shorthand for <math>\exists A \in \mathcal{T}_C. [x : A \in \Gamma]</math>, likewise, if <math>x</math> is not typed in the context we use <math>x \notin \Gamma</math>.</li> </ul>	
For example:	
$\Gamma_{\text{my context}} = \{x : A, y : B, c : B\}$	

$\rightarrow$  is used for function types, it is right associative, so:

$$(A \rightarrow B) \rightarrow C \rightarrow D \equiv (A \rightarrow B) \rightarrow (C \rightarrow D)$$

### 3.0.1 Curry Type Assignment

$$(Ax) : \frac{}{\Gamma, x : A \vdash_C x : A} \quad (\rightarrow I) : \frac{\Gamma, x : A \vdash_C M : B}{\Gamma \vdash_C \lambda x.M : A \rightarrow B} (x \notin \Gamma) \quad (\rightarrow E) : \frac{\Gamma \vdash_C M_1 : A \rightarrow B \quad \Gamma \vdash_C M_2 : A}{\Gamma \vdash_C M_1 M_2 : B}$$

- We can extend barendregt's convention to ommit the side-condition on  $\rightarrow I$  by adding the assertion that:

$$\Gamma \vdash M : A \text{ we ensure } \forall x \in bv(M). [x \notin \Gamma]$$

- The definition provided is *sound*:

$$(\Gamma \vdash_c M : A) \wedge (M \rightarrow_\beta^* N) \Rightarrow \Gamma \vdash_C N : A$$

Some terms are not typeable under this definition, as self-application is not possible:

- $\lambda x.x x$  is not typeable, neither is  $\lambda f.(\lambda x.f(x x))(\lambda x.f(x x))$
- Type assignment rules do not cover approximants, and hence they are not typeable.

### Self Application

### Example Question 3.0.1

Is it possible to type *self-application*  $x x$ ?

We can attempt to use the inference system, however run into a contradiction:

$$\frac{\frac{}{\Gamma, x : A \rightarrow B \vdash_C x : A \rightarrow B}^{(Ax)} \quad \frac{}{\Gamma, x : A \vdash_C x : A}^{(Ax)}}{\Gamma, x : ? \vdash_C x x : B}(\rightarrow E)$$

Hence we need a type such that  $A \rightarrow B = A$ .

## 3.0.2 Important Lemmas For Type Assignment

### Term Substitution

$$\exists C.[(\Gamma, x : C \vdash_C M : A) \wedge (\Gamma \vdash_C N : C)] \Rightarrow \Gamma \vdash_C M[N/x] : A$$

### Free Variables

$$\Gamma \vdash_C M : A \wedge x \in fv(M) \Rightarrow \exists B \in \mathcal{T}_C.[x : B \in \Gamma]$$

All free variables in  $M$  are typed.

### Weakening

$$\Gamma \vdash_C M : A \wedge \Gamma' \text{ is such that } \forall x : B \in \Gamma'. [x : B \in \Gamma \vee (x \notin fv(M) \wedge x \notin bv(M))] \Rightarrow \Gamma' \vdash_C M : A$$

We can create a new context  $\Gamma'$  that types variables  $x, y, z, \dots$ . If for every variable in the context  $\Gamma'$  it is either not in  $\Gamma$ , or is in  $\Gamma$  with the same type, then we can use  $\Gamma'$  to type  $M$ .

### Thinning

$$\Gamma, x : B \vdash_C M : A \wedge x \notin fv(M) \Rightarrow \Gamma \vdash_C M : A$$

If a variable is not free in  $M$ , then we do not need a type for it.

## 3.1 Principle Type Property

Principle type theory expresses the idea that a whole family of types could be assigned to a term, however only one is the *principle type*.

### Type Substitution

### Definition 3.1.1

$$(\varphi \mapsto C) : \mathcal{T}_C \rightarrow \mathcal{T}_C \text{ where } \varphi \text{ is a type variable and } C \in \mathcal{T}_C$$

Substitution is defined by:

$$\begin{aligned} (\varphi \mapsto C) \quad \varphi &= C \\ (\varphi \mapsto C) \quad \varphi' &= \varphi' & (\varphi \neq \varphi') \\ (\varphi \mapsto C) \quad A \rightarrow B &= ((\varphi \mapsto C) A) \rightarrow ((\varphi \mapsto C) B) \end{aligned}$$

Here  $(\varphi \mapsto C)$  is a substitution substituting the type variable  $\varphi$  for the type  $C$

$$S_1 \circ S_2 \text{ means } S_1 \circ S_2 A = S_1(S_2 A) \quad S \Gamma = \{x : S B \mid x : B \in \Gamma\} \quad S(\Gamma; A) = \langle S \Gamma; S A \rangle$$

- If there is a substitution  $S$  such that  $S A = B$  then  $B$  is the *substitution instance* of  $A$ .
- $Id_S$  (*identity substitution*) maps every type variable to itself.

For each typeable term  $M$  there is a principal pair:

$\langle \Pi; P \rangle$  where  $\Pi$  is a context and  $P \in \mathcal{T}_C$  such that  $\forall \Gamma, A \in \mathcal{T}_C. \exists$  substitution  $S. [S \langle \Pi; P \rangle = \langle \Gamma; A \rangle]$

Soundness	Definition 3.1.2	Completeness	Definition 3.1.3
A logical system is sound if every formula provable using the system is logically valid according to the semantics of the system.		A logical system is complete if any true statement can be proved using the system.	
$Provable \Rightarrow True$		$True \Rightarrow Provable$	

This definition is sound, for every substitution  $S$ :

if there is a derivation for  $\Gamma \vdash_C M : A$  then we can construct a derivation for  $S \Gamma \vdash_C M : S A$

### 3.1.1 Unification

Robinson's Unification	Definition 3.1.4
$ \begin{aligned} unify \quad \varphi \quad \varphi &= (\varphi \mapsto \varphi) \\ unify \quad \varphi \quad B &= (\varphi \mapsto B) \text{ given } \varphi \text{ does not occur in } B \\ unify \quad A \quad \varphi &= unify \quad \varphi \quad A \\ unify \quad (A \rightarrow B) \quad (C \rightarrow D) &= S_1 \circ S_2 \text{ where} \\ &\quad S_1 = unify \quad A \quad C \\ &\quad S_2 = unify \quad (S_1 B) \quad (S_1 D) \end{aligned} $ <ul style="list-style-type: none"> <li>• Unification is associative and commutative</li> <li>• It returns the most general unifier of two types (the <i>common substitution instance</i>)</li> </ul>	

Robinson's Unification can be generalised to unify contexts.

$$\begin{aligned}
unifyContexts \quad (\Gamma_1, x : A) \quad \Gamma_2 &= unifyContexts \quad \Gamma_1 \quad \Gamma_2 \text{ given } x \text{ does not occur in } \Gamma_2 \\
unifyContexts \quad \emptyset \quad \Gamma_2 &= Id_S \\
unifyContexts \quad (\Gamma_1, x : A) \quad (\Gamma_2, x : B) &= S_1 \circ S_2 \text{ where} \\
&\quad S_1 = unify \quad A \quad B \\
&\quad S_2 = unifyContexts \quad (S_1 \Gamma_1) \quad (S_1 \Gamma_2)
\end{aligned}$$

### 3.1.2 Curry Principle Pair

Curry Principle Pair	Definition 3.1.5
Every term $M$ has a (Curry) Principle Pair defined as $pp_c M = \langle \Pi; P \rangle$ by:	
$pp_c \quad x \quad = \langle x : \varphi; \varphi \rangle \text{ where } \varphi \text{ is fresh}$	
$ pp_c \quad \lambda x. M \quad = \begin{cases} \langle \Pi'; A \rightarrow P \rangle & (\Pi = \Pi', x : A) \\ \langle \Pi; \varphi \rightarrow P \rangle & (x \notin \Pi) \end{cases} $ <p style="text-align: center;">where <math>\langle \Pi; P \rangle = pp_c M</math>  <math>\varphi</math> is fresh</p>	
$ pp_c \quad M \quad N \quad = S_2 \circ S_1 \langle \Pi_1 \cup \Pi_2; \varphi \rangle $ <p style="text-align: center;">where <math>\langle \Pi_1; P_1 \rangle = pp_c M</math>  <math>\langle \Pi_2; P_2 \rangle = pp_c N</math>  <math>S_1 = unify \quad P_1 \quad (P_2 \rightarrow \varphi)</math>  <math>S_2 = unifyContexts \quad (S_1 \Pi_1) \quad (S_1 \Pi_2)</math>  <math>\varphi</math> is fresh</p>	

Substitution is complete:

$$\forall \Gamma, M \in \Lambda, A \in \mathcal{T}_c. [\Gamma \vdash_c M : A \Rightarrow \exists \Pi, P \in \mathcal{T}_c. S. [pp_c M = \langle \Pi; P \rangle \wedge s \Pi \subseteq \Gamma \wedge S P = A]]$$

## Chapter 4

# Polymorphism

We can extend the  $\lambda$ -calculus to allow for functions that are *polymorphic* (can be applied to many different types of inputs).

- We can extend to include names and definitions (e.g  $name = M$ )
- When type checking we can associate a call to a name with its definition, avoiding the need to re-type check for each call to a function.

### 4.1 Language $\Lambda^N$

$\Lambda^N$  is Lambda Calculus with names. The syntax is as follows:

$$\begin{aligned} name &::= \text{'A string of characters'} \\ N, M &::= x \mid name \mid \lambda x. N \mid M N \\ Defs &::= Defs; name = M \mid \epsilon \text{ where } M \text{ is closed and name-free} \\ Program &::= Defs : M \end{aligned}$$

Reduction on terms can be defined by an inference system.

$$\begin{array}{c} \frac{}{(\lambda x. M) N \rightarrow M[N/x]} \\ \text{Substitution of terms} \end{array} \qquad \frac{}{name \rightarrow M}^{(name = M \in Defs)} \\ \text{Substituting of names for definitions (inlining)} \\[10pt] \frac{M \rightarrow N}{\lambda x. M \rightarrow \lambda x. N} \quad \frac{M \rightarrow N}{M P \rightarrow N P} \quad \frac{M \rightarrow N}{P M \rightarrow P N} \\ \text{Reduction of terms} \\[10pt] \frac{M \rightarrow N}{M \rightarrow^* N} \quad \frac{}{M \rightarrow^* M} \quad \frac{M \rightarrow^* N \quad N \rightarrow^* P}{M \rightarrow^* P} \\ \text{Transitive closure of reduction} \\[10pt] \frac{M \rightarrow N}{Defs : M \rightarrow Defs : N} \\ \text{Reduction on Programs} \end{array}$$

- Names are closed  $\lambda$ -terms (have no free variables).
- If a name is used but not defined, then the program is irreducible.
- Programs written in  $\Lambda^N$  can be translated to  $\Lambda$  by substituting names.

We can translate using the transformation  $\langle \cdot \rangle_\lambda : \Lambda^N \rightarrow \Lambda$ :

$$\begin{aligned} \langle x \rangle_\lambda &= x \\ \langle name \rangle_\lambda &= \begin{cases} \langle M \rangle_\lambda & \text{if } (name = M) \in Defs \\ undefined & \text{otherwise} \end{cases} \\ \langle \lambda x. N \rangle_\lambda &= \lambda x. \langle N \rangle_\lambda \\ \langle \lambda N M \rangle_\lambda &= \langle \lambda N \rangle_\lambda \langle \lambda M \rangle_\lambda \end{aligned}$$

## 4.2 Type Assignment for $\Lambda^N$

By extending Curry's type assignment system for  $\lambda$ -calculus we must consider the types of names in definitions.

Environment	Definition 4.2.1
<p>An environment <math>\varepsilon</math> is a mapping on <math>names \rightarrow \mathcal{T}_c</math>.</p> <ul style="list-style-type: none"> <li>• Similar to a context, but for names rather than terms.</li> <li>• <math>\varepsilon, name : A = \varepsilon \cup \{name : A\}</math> where either <math>name : A \in \varepsilon</math> or <math>name</math> does not occur in <math>\varepsilon</math>.</li> </ul>	

$$(Ax) : \frac{}{\Gamma, x : A; \varepsilon \vdash x : A} \quad (\rightarrow I) : \frac{\Gamma, x : A; \varepsilon \vdash N : B}{\Gamma; \varepsilon \vdash \lambda x. N : A \rightarrow B} \quad (\rightarrow E) : \frac{\Gamma; \varepsilon \vdash P : A \rightarrow B \quad \Gamma; \varepsilon \vdash Q : A}{\Gamma; \varepsilon \vdash P Q : B}$$

We have extended the curry type inference to include the environment  $\varepsilon$ .

$$(\epsilon) : \frac{}{\varepsilon \vdash \epsilon : \Diamond}$$

We do not need to consider contexts (definitions use closed terms, no free variables from a context are required to type).  $\Diamond$  is not a type, but rather notation of showing there is a type.

$$(Defs) : \frac{\varepsilon \vdash Defs : \Diamond \quad \emptyset; \emptyset \vdash M : A}{\varepsilon, name : A \vdash Defs; name = M : \Diamond}$$

A name can be defined, it must be closed (hence why context is  $\emptyset$ ). Notice this definition ensures definitions are closed and name-free as the rule provides an empty environment and context.

$$(Call) : \frac{}{\Gamma; \varepsilon, name : A \vdash name : S A}$$

$$(Program) : \frac{\varepsilon \vdash Defs : \Diamond \quad \Gamma; \varepsilon \vdash M : A}{\Gamma; \varepsilon \vdash Defs : M : A}$$

## 4.3 Principal Types for $\Lambda^N$

$$\begin{aligned}
pp_{\Lambda^N} x \quad \varepsilon &= \langle x : \varphi; \varphi \rangle \text{ where } \varphi \text{ is fresh} \\
pp_{\Lambda^N} name \quad \varepsilon &= \langle \emptyset; FreshInstance(\varepsilon name) \rangle \\
pp_{\Lambda^N} (\lambda x. M) \quad \varepsilon &= \begin{cases} \langle \Pi'; A \rightarrow P \rangle & (\Pi = \Pi', x : A) \\ \langle \Pi; \varphi \rightarrow P \rangle & (x \notin \Pi) \end{cases} \\
&\quad \text{where } \langle \Pi; P \rangle = pp_{\Lambda^N} M \quad \varepsilon \\
&\quad \varphi \text{ is fresh} \\
pp_{\Lambda^N} (M N) \quad \varepsilon &= S_2 \circ S_1 \langle \Pi_1 \cup \Pi_2; \varphi \rangle \\
&\quad \text{where } \langle \Pi_1; P_1 \rangle = pp_{\Lambda^N} M \quad \varepsilon \\
&\quad \langle \Pi_2; P_2 \rangle = pp_{\Lambda^N} N \quad \varepsilon \\
&\quad S_1 = unify P_1 P_2 \rightarrow \varphi \\
&\quad S_2 = unifyContexts (S_1 \Pi_1) (S_1 \Pi_2) \\
&\quad \varphi \text{ is fresh}
\end{aligned}$$

We also need to define  $pp_{\Lambda^N}$  for definitions.

$$\begin{aligned}
BuildEnv (Defs; name = M) &= (BuildEnv Defs), name : A \text{ where } \langle \emptyset; A \rangle = pp_{\Lambda^N} M \quad \emptyset \\
BuildEnv \epsilon &= \emptyset
\end{aligned}$$

Hence we can now define:

$$pp_{\Lambda^N} (Defs; M) = pp_{\Lambda^N} M \quad \varepsilon \text{ where } \varepsilon = BuildEnv Defs$$

- For each name encountered, the environment is checked to find its principle type, a fresh instance of this type is taken (with all type variables replaced by fresh ones) this allows for polymorphism.

Derive the  $\Lambda^N$  type for  $\lambda x.x$  where it is named  $Id$

$$\frac{\frac{\overline{\emptyset \vdash \epsilon : \Diamond}}{(\epsilon)} \quad \frac{\overline{x : \varphi; \emptyset \vdash x : \varphi}^{(Ax)} \quad \overline{\emptyset; \emptyset \vdash_c \lambda x.x : \varphi \rightarrow \varphi}^{(\rightarrow I)}}{\overline{Id : \varphi \rightarrow \varphi \vdash I = \lambda x.x}^{(Defs)}} \quad \frac{\overline{\emptyset; Id : \varphi \rightarrow \varphi \vdash Id \quad Id : A \rightarrow A}^{(Program)} \quad \overline{Call_1 \quad Call_2}^{(\rightarrow E)}}{\overline{\emptyset; Id : \varphi \rightarrow \varphi \vdash I = \lambda x.x : Id \quad Id : A \rightarrow A}^{(Program)}}$$

$$Call_1 = \overline{\emptyset; I : \varphi \rightarrow \varphi \vdash I : (A \rightarrow A) \rightarrow A \rightarrow A}^{(Call)}$$

$$Call_2 = \overline{\emptyset; I : \varphi \rightarrow \varphi \vdash I : A \rightarrow A}^{(Call)}$$



# Chapter 5

## Recursion

We can extend  $\Lambda^N$  to include recursion as language  $\Lambda^{NR}$ .

- Definitions can reference their own names, as well as other's names (e.g for mutually recursive functions)

### 5.1 Language $\Lambda^{NR}$

$name ::= \text{'A string of characters'}$   
 $N, M ::= x \mid name \mid \lambda x. N \mid M \ N$   
 $Defs ::= Defs; name = M \mid Defs; (rec \ name = M) \mid \epsilon$  where  $M$  is closed  
 $Program ::= Defs : M$

The requirement that  $M$  be name-free is removed, and a function labeled  $rec$  can be recursive.

Y combinator	Definition 5.1.1
Can be used to encode recursion:	$Y = \lambda f. (\lambda x. f x x) (\lambda x. f x x)$ $F = C[F] \rightarrow Y (\lambda f. C[f])$

Factorial	Example Question 5.1.1
Write factorial in $\Lambda^{NR}$ given you can use arithmetic and the $Cond$ function. Then encode it using the $Y$ combinator.	
$Factorial = \lambda n. (Cond \ (n == 0) \ 1 \ (n \times (Factorial \ (n - 1))))$	
And with the $Y$ combinator:	
$Fac = Y. (\lambda f n. Cond \ (n == 0) \ 1 \ (n \times f(n - 1)))$	

We cannot directly translate  $\Lambda^{NR}$  to lambda calculus as with  $\Lambda^N$ , and instead must alter recursive functions to make use of the  $Y$  combinator.

$Y$  is not typeable under the  $\vdash_c$  scheme discussed in these notes, so we must add an extension:

$$M, N ::= \dots \mid Y$$

$$Y \ M \rightarrow M(Y \ M)$$

$$\overline{\Gamma \vdash Y : (A \rightarrow A) \rightarrow A}$$

Add  $Y$  as a special term in the syntax. Add the reduction rule for  $Y$ .

Add a type assignment rule.

### 5.2 Type Assignment for $\Lambda^{NR}$

$$(Ax) : \frac{}{\Gamma, x : A; \varepsilon \vdash x : A} \quad (\rightarrow I) : \frac{\Gamma, x : A; \varepsilon \vdash N : B}{\Gamma; \varepsilon \vdash \lambda x. N : A \rightarrow B} \quad (\rightarrow E) : \frac{\Gamma; \varepsilon \vdash P : A \rightarrow B \quad \Gamma; \varepsilon \vdash Q : A}{\Gamma; \varepsilon \vdash P \ Q : B}$$

The main 3 typing rules remain unchanged.

$$(Call) : \frac{}{\Gamma; \varepsilon, name : A \vdash name : S \ A} \quad (Rec \ Call) : \frac{}{\Gamma; \varepsilon, rec \ name : A \vdash name : A}$$

- Call remains the same (still just substitutes the definition)
- A recursive call is added, however the type cannot be substituted for this as the definition internally relies on the type.

$$(Def) : \frac{\varepsilon \vdash Defs : \Diamond \quad \emptyset; \varepsilon \vdash M : A}{\varepsilon, name : A \vdash Defs; name = M : \Diamond} \quad (Rec \ Def) : \frac{\varepsilon \vdash Defs : \Diamond \quad \emptyset; \varepsilon, rec \ name : A \vdash M : A}{\varepsilon, name : A \vdash Defs; rec \ name = M : \Diamond}$$

- Definitions are no longer name-free and thus we must carry the environment in *Def*.
- Recursive calls are typed with the same environment.

$$(\epsilon) : \frac{}{\varepsilon \vdash \epsilon : \Diamond} \quad (Program) : \frac{\varepsilon \vdash Defs : \Diamond \quad \Gamma; \varepsilon \vdash M : A}{\Gamma; \varepsilon \vdash Defs : M : A}$$

### 5.3 Principle Types for $\Lambda^{NR}$

$$\begin{aligned} pp_{\Lambda^{NR}} \ x \quad \varepsilon &= \langle x : \varphi; \varphi; \varepsilon \rangle \text{ where } \varphi \text{ is fresh} \\ pp_{\Lambda^{NR}} \ name \quad \varepsilon &= \begin{cases} \langle \emptyset; A; \varepsilon \rangle & (rec \ name : A \in \varepsilon) \\ \langle \emptyset; FreshInstance(\varepsilon \ name); \varepsilon \rangle & (name : A \in \varepsilon) \end{cases} \\ pp_{\Lambda^{NR}} \ (\lambda x.M) \quad \varepsilon &= \begin{cases} \langle \Pi'; A \rightarrow P; \varepsilon' \rangle & (\Pi = \Pi', x : A) \\ \langle \Pi; \varphi \rightarrow P; \varepsilon' \rangle & (x \notin \Pi) \end{cases} \\ &\quad \text{where } \langle \Pi; P; \varepsilon' \rangle = pp_{\Lambda^{NR}} \ M \ \varepsilon \\ &\quad \varphi \text{ is fresh} \\ pp_{\Lambda^{NR}} \ (M \ N) \quad \varepsilon &= S_2 \circ S_1 \langle \Pi_1 \cup \Pi_2; \varphi; \varepsilon'' \rangle \\ &\quad \text{where } \begin{aligned} \langle \Pi_1; P_1; \varepsilon' \rangle &= pp_{\Lambda^{NR}} \ M \ \varepsilon \\ \langle \Pi_2; P_2; \varepsilon'' \rangle &= pp_{\Lambda^{NR}} \ N \ \varepsilon \\ S_1 &= unify \ P_1 \ P_2 \rightarrow \varphi \\ S_2 &= unifyContexts \ (S_1 \Pi_1) \ (S_1 \Pi_2) \\ \varphi &\text{ is fresh} \end{aligned} \end{aligned}$$

For defs we must modify the *buildEnv* function to use environments:

$$\begin{aligned} BuildEnv \ (defs; name = M) \ \varepsilon &= (BuildEnv \ Defs \ \varepsilon), name : A \text{ where } \langle \emptyset; A; \varepsilon \rangle = pp_{\Lambda^{NR}} \ M \ \varepsilon \\ BuildEnv \ (defs; rec \ name = M) \ \varepsilon &= (BuildEnv \ Defs \ \varepsilon), name : S \ A \\ &\quad \text{where } \begin{aligned} \langle \emptyset; A; \varepsilon' \rangle &= pp_{\Lambda^{NR}} \ M(\varepsilon, rec \ name : \varphi) \\ S &= unify \ A \ B \\ rec \ name : B &\in \varepsilon' \\ \varphi &\text{ is fresh} \end{aligned} \\ BuildEnv \ \epsilon \ \varepsilon &= \varepsilon \\ pp_{\Lambda^{NR}}(Defs; M) & \end{aligned}$$

Hence we can now define  $pp_{\Lambda^{NR}}$  for *Defs*:

$$pp_{\Lambda^{NR}} \ (Defs; M) = pp_{\Lambda^{NR}} \ M \ \varepsilon \text{ where } \varepsilon = BuildEnv \ Defs \ \emptyset$$

# Chapter 6

## Milner's ML

$\mathcal{L}_{ML}$

Definition 6.0.1

$\mathcal{L}_{ML}$  is a simple programming language supporting *shallow polymorphic* procedures on a wide variety of objects.

- It is an extension of  $\lambda$ -calculus
- Adds a construct for expressing recursion
- Adds a construct for expressing sub-terms can be used in different ways.

A new type-assignment algorithm is paired with  $\mathcal{L}_{ML}$  called  $\mathcal{W}$ :

- *Semantically Sound* - all typed programs are correct.
- *Syntactically Sound* - if  $\mathcal{W}$  accepts a program, then it is well-typed.

### 6.1 The ML Type Assignment System

ML expressions are of the form:

$$E ::= x \mid c \mid \lambda x.E \mid E_1 E_2 \mid \text{let } x = E_1 \text{ in } E_2 \mid \text{fix } g.E$$

where:

- $x$  is bound over  $E_2$  in  $\text{let } x = E_1 \text{ in } E_2$
- $g$  is bound over  $E$  in  $\text{fix } g.E$
- $c$  is a term constant, such as a number, character or operator

#### 6.1.1 Term Substitution

Term substitution is defined as with the following rules:

$$x[E/x] = x$$

$$y[E/x] = y \quad (y \neq x)$$

Basic substitution of variables.

$$(\lambda y.E')[E/x] = \lambda y.(E'[E/x])$$

$$(E_1 E_2)[E/x] = E_1[E/x] E_2[E/x]$$

Substitution of sub-terms.

$$(\text{let } x = E_1 \text{ in } E_2)[E/x] = \text{let } y = E_1[E/x] \text{ in } E_2[E/x]$$

$$(\text{fix } g.E')[E/x] = \text{fix } g.E'[E/x]$$

let statements and fixed point (for recursion)

Note that *barendregt's convention* assumed here.

- The *let* construction is added to cover cases where  $(\lambda x.E_1)E_2$  is not typeable but where the contraction  $E_1[E_2/x]$  is typeable.
- The *fix* construction introduces model recursion. It is not a combinator, but rather another abstraction mechanism (e.g like  $\lambda$ ).

### 6.1.2 Reduction

Reduction on  $\mathcal{LML}$  is  $\rightarrow_{ML}$  and is defined as an extension of  $\rightarrow_\beta$ , with the additional rules:

$$\begin{aligned} \text{let } x_1 = E_1 \text{ in } E_2 &\rightarrow_{ML} E_2[E_1/x] \\ \text{fix } g.E &\rightarrow_{ML} E[(\text{fix } g.E)/g] \end{aligned}$$

We also add some contextual rules.

$$E \rightarrow_{ML} E' \Rightarrow \begin{cases} \text{let } x = E \text{ in } E_2 &\rightarrow_{ML} \text{let } x = E' \text{ in } E_2 \\ \text{let } x = E_1 \text{ in } E &\rightarrow_{ML} \text{let } x = E_1 \text{ in } E' \\ \text{fix } g.E &\rightarrow_{ML} \text{fix } g.E' \end{cases}$$

Under reduction both  $\text{let } x = E_2 \text{ in } E_1$  and  $(\lambda x.E_1) E_2$  are *reducible expressions* and both reduce to  $E_1[E_2/x]$

- $(\lambda x.E_1) E_2$  semantically interpreted as a function with an operand  $x$
- $\text{let } x = E_2 \text{ in } E_1$  interpreted as a substitution.

Type assignment treats both differently.

### 6.1.3 Type Assignment

The set of types is defined similarly to with curry types ( $\mathcal{T}_c$ ).

- Extended with type constants ' $C$ ' that includes  $\text{int}, \text{bool}, \dots$
- Ranged over by type  $A, B, \dots$  much like with  $\mathcal{T}_c$ .
- Types can be *quantified*, creating *generic types* / *type schemes* ranged over by  $\sigma, \tau, \dots$

$$\begin{aligned} A, B &::= \varphi \mid c \mid (A \rightarrow B) && \text{(basic types)} \\ \sigma, \tau &::= A \mid (\forall \varphi. \tau) && \text{(polymorphic types)} \end{aligned}$$

Types of the form  $\forall \varphi. \tau$  are called *quantified types*.

- $(\forall \varphi_1. (\forall \varphi_2. \dots (\forall \varphi_n. A) \dots))$  is abbreviated by  $\forall \vec{\varphi}. A$
- $\varphi$  is bound in  $\forall \varphi. \tau$
- Free and bound type variables can be defined just as with variables in  $\lambda$ -calculus, but must have names kept separate.

*ML type substitution* is defined as:

$$\begin{aligned} (\varphi \mapsto C) \quad \varphi &= C \\ (\varphi \mapsto C) \quad c &= c \\ (\varphi \mapsto C) \quad \varphi' &= \varphi' && (\varphi' \neq \varphi) \\ (\varphi \mapsto C) \quad A \rightarrow B &= ((\varphi \mapsto C)A) \rightarrow ((\varphi \mapsto C)B) \end{aligned} \qquad \begin{aligned} (\varphi \mapsto C) \quad \forall \varphi'. \psi &= \forall \varphi'. ((\varphi \mapsto C)\psi) \end{aligned}$$

Basic type substitutions

Quantified types

Unification is also extended with type constants as:

$$\begin{aligned} \text{unify} \quad \varphi \quad c &= (\varphi \mapsto c) \\ \text{unify} \quad c \quad \varphi &= \text{unify} \quad \varphi \quad c \\ \text{unify} \quad c \quad c &= Id_S \end{aligned}$$

- Here a unification of all other cases including a type constant will fail (e.g cannot unify  $\text{int}$  and  $\text{bool}$ )
- Types are considered modulo a kind of  $\alpha$ -conversion (similar to barendregt's convention - avoid type variable capture)
- As  $\varphi'$  is bound in  $\forall \varphi'. \psi$  we can assume in  $(\varphi \mapsto C) \forall \varphi'. \psi$  we have  $\varphi \neq \varphi'$  and  $\varphi' \notin fv(C)$ .
- As we can have free type variables, the set of types occurring in  $\forall \varphi_1 \dots \forall \varphi_n. A$  is not necessarily  $\{\varphi_1, \dots, \varphi_n\}$ .

$$\bar{\Gamma} A = \forall \vec{\varphi}. A$$

$\vec{\varphi}$  appear free in  $A$ , but are not in the context of  $A$ .

For the inference system expressing type assignment, we include a function  $\nu$  which maps constants to their type (e.g a constant type such as Char, Int or a closed polymorphic type).

$$\begin{array}{ll}
(Ax) : \frac{}{\Gamma, x : \tau \vdash x : \tau} & (C) : \frac{}{\Gamma \vdash c : \nu c} \\
\text{Basic substitution of free variable} & \text{Substituting constants} \\
\\
(\rightarrow I) : \frac{\Gamma, x : A \vdash E : B}{\Gamma \vdash \lambda x. E : A \rightarrow B} & (\rightarrow E) : \frac{\Gamma \vdash E_1 : A \rightarrow B \quad \Gamma \vdash E_2 : A}{\Gamma \vdash E_1 E_2 : B} \\
\\
(\text{let}) : \frac{\Gamma \vdash E_1 : \tau \quad \Gamma, x : \tau \vdash E_2 : B}{\Gamma \vdash \text{let } x = E_1 \text{ in } E_2 : B} & (\text{fix}) : \frac{\Gamma, g : A \vdash E : A}{\Gamma \vdash \text{fix } g. E : A} \\
\\
(\forall I) : \frac{\Gamma \vdash E : \tau}{\Gamma \vdash E : \forall \varphi. \tau} (\varphi \text{ not free in } \Gamma) & (\forall E) : \frac{\Gamma \vdash E : \forall \varphi. \tau}{\Gamma \vdash E : \tau[A/\varphi]}
\end{array}$$

Quantification is introduced to model substitution operations on types, rather than replacing all type variables at once.

- $\forall \varphi : \tau$  - all occurrences of type variable  $\varphi$  can be replaced by some basic type.
- The side condition on  $\forall I$  ensures that the type variables used do not also occur in the context (there is no reference back to the context).

We can model the substitution of  $\varphi$  in  $A$ , by type  $B$  as  $(\varphi \mapsto B) A$ .

$$\frac{\frac{\emptyset \vdash_{ML} E : A}{\emptyset \vdash_{ML} E : \forall \varphi. A} (\forall I)}{\emptyset \vdash_{ML} E : A[B/\varphi]} (\forall E)$$

The let construct corresponds to definitions in  $A^{NR}$ .

- Can occur anywhere within a term.
- Given  $\text{let } x = E_1 \text{ in } E_2$ ,  $E_1$  does not need to be a *closed-term*, so it is possible to define terms that are *partially-polymorphic* (a term of type  $\forall \vec{\varphi}. A$  where  $A$  contains free type variables).
- When applying  $\forall I$  only the type variable that we attempt to bind must not occur in the context.

To allow for recursion to be typed, the syntax for fix is added.

- Previously we have seen **Y** added as a typed constant.
- It is also possible to solve this by defining *letrec* as a combination of *let* and *fix* ( $\text{letrec } g = \lambda x. E_1 \text{ in } E_2$ )

## 6.1.4 Lemmas for Type Assignment

### Free Variables

$$(\Gamma \vdash_{ML} E : \tau \wedge x \in \text{fv}(E)) \Rightarrow \exists \sigma. [x : \sigma \in \Gamma]$$

All free variables in some expression must have a type in the context.

### Weakening

$$(\Gamma \vdash_{ML} E : \tau \wedge \forall x : \sigma \in \Gamma' [x : \sigma \in \Gamma \vee x \notin (\text{fv}(E) \cup \text{bv}(E))]) \Rightarrow \Gamma' \vdash_{ML} E : \tau$$

Given some context  $\Gamma$ , any context that extends  $\Gamma$  without adding any of  $E$ 's variables is equivalent.

### Thinning

$$(\Gamma x : \sigma \vdash_{ML} E : \tau \wedge x \notin \text{fv}(E)) \Rightarrow \Gamma \vdash_{ML} E : \tau$$

We can remove variables that are not free in  $E$  from the context, and the context will still be able to type  $E$ .

$>_{\Gamma}$ 

Definition 6.1.1

$$\Gamma \vdash_{ML} E : A[B/\varphi]$$

The smallest reflexive and transitive relation such that:

$$\begin{aligned} \rho &>_{\Gamma} \forall \varphi. \rho && (\varphi \text{ is not free in } \Gamma \text{ and not bound in } \rho) \\ \forall \varphi. \rho &>_{\Gamma} \rho[B/\varphi] \end{aligned}$$

Where there are no free  $\varphi'$  in  $A$ .

- If  $\sigma >_{\Gamma} \tau$  then  $\tau$  is a *generic instance* of  $\sigma$ .
- Each context  $\Gamma$  induces a new relation.
- This relation represents applying the  $\forall I$  and  $\forall E$  steps.

$$\Gamma \vdash_{ML} E : \sigma \wedge \sigma >_{\Gamma} \tau \Rightarrow \Gamma_{ML} E : \tau$$

- (1)  $\Gamma \vdash_{ML} x : \sigma \Rightarrow \exists x : \tau \in \Gamma. \tau >_{\Gamma} \sigma$
- (2)  $\Gamma \vdash_{ML} \lambda x. E : \sigma \Rightarrow \exists A, B. \begin{aligned} &\Gamma, x : A \vdash_{ML} E : B \\ &\wedge \sigma = \forall \vec{\varphi}_i. A \rightarrow B \\ &\wedge A \rightarrow B >_{\Gamma} \sigma \end{aligned}$
- (3)  $\Gamma \vdash_{ML} E_1 E_2 : \sigma \Rightarrow \exists A, B. \begin{aligned} &\Gamma \vdash_{ML} E_1 : A \rightarrow B \\ &\wedge \Gamma \vdash_{ML} E_2 : A \\ &\wedge B >_{\Gamma} \sigma \end{aligned}$
- (4)  $\Gamma \vdash_{ML} \text{fix } g. E : \sigma \Rightarrow \exists A. \begin{aligned} &\Gamma, g : A \vdash_{ML} E : A \\ &\wedge \sigma = \sigma = \forall \vec{\varphi}_i. A \\ &\wedge A >_{\Gamma} \sigma \end{aligned}$
- (5)  $\Gamma \vdash_{ML} \text{let } x = E_1 \text{ in } E_2 : \sigma \Rightarrow \exists A, \tau \begin{aligned} &\Gamma, x : \tau \vdash_{ML} E_2 : A \\ &\wedge \Gamma \vdash_{ML} E_1 : \tau \\ &\wedge A >_{\Gamma} \sigma \end{aligned}$

**System F***Extra Fun!* 6.1.1

The ML type assignment is a restriction on the *polymorphic type discipline* (System F).

- In the ML type assignment covered in these notes,  $\forall$  occurs outside of a type (*shallow polymorphism*). It is also decidable.
- In System F  $\forall$  is a general type constructor, so  $A \rightarrow \forall \varphi. B$  is a valid type. It is not decidable.

**Complex Types**

Example Question 6.1.1

Type let  $i = \lambda x. x$  in  $i i$ .

$$\frac{\frac{\frac{}{(Ax)}{x : \varphi \vdash x : \varphi}(\rightarrow I)}{\emptyset \vdash \lambda x. x : \varphi \rightarrow \varphi}(\forall I)}{\emptyset \vdash \lambda x. x : \forall \varphi. \varphi \rightarrow \varphi} \quad \frac{\frac{(1) \quad (2)}{i : \forall \varphi. \varphi \rightarrow \varphi \vdash i i : A \rightarrow A}(\rightarrow E)}{\emptyset \vdash \text{let } i = \lambda x. x \text{ in } i i : A \rightarrow A}(\text{let})$$

where:

$$(1) = \frac{\frac{}{(Ax)}{i : \forall \varphi. \varphi \rightarrow \varphi \vdash i : \forall \varphi. \varphi \rightarrow \varphi}}{i : \forall \varphi. \varphi \rightarrow \varphi \vdash i : (A \rightarrow A) \rightarrow A \rightarrow A}(\forall E)$$

$$(2) = \frac{\overline{i : \forall \varphi. \varphi \rightarrow \varphi \vdash i : \forall \varphi. \varphi \rightarrow \varphi}^{(Ax)}}{i : \forall \varphi. \varphi \rightarrow \varphi \vdash i : A \rightarrow A} (\forall E)$$

## Addition

## Example Question 6.1.2

Express Addition in ML.

We can use type constants by defining then in  $\nu$ :

$$\nu x = \begin{cases} Num \rightarrow Num & x = Succ \\ Num \rightarrow Num & x = Pred \\ Num \rightarrow Bool & x = IsZero \\ \forall \varphi. Bool \rightarrow \varphi \rightarrow \varphi \rightarrow \varphi & x = Cond \\ \vdots & \vdots \end{cases}$$

We can define it recursively as:

$$Add = \lambda xy. Cond (IsZero x) y (Succ (Add (Pred x) y))$$

This recursion can be implemented in ML using fix.

$$Add = \text{fix } a. \lambda xy. Cond (IsZero x) y (Succ (a (Pred x) y))$$

## 6.2 Milner's $\mathcal{W}$

### Milner's Type Assignment Algorithm

### Definition 6.2.1

Milner's  $\mathcal{W}$  is a type assignment algorithm for ML.

- It has a *principle type property* - given any  $\Gamma$  and  $E$  there is a principle type computed by  $\mathcal{W}$ .
- Its does not have the *principle pair property* as if  $\Gamma, x : \tau \vdash_{ML} E : A$  may exist, but  $\lambda x. E$  may not be typeable.
- Type assignment is decidable.

It is complete, given some  $E$ , contexts  $\Gamma$  and  $\Gamma'$  and type  $A$ :

$$\Gamma' \text{ is an instance of } \Gamma \wedge \Gamma' \vdash_{ML} E : A \Rightarrow \mathcal{W} \Gamma E = \langle S, B \rangle \wedge \exists S'. [\Gamma' = S'(S \Gamma) \wedge S'(S B) >_{\Gamma'} A]$$

It is also sound:

$$\forall E. [\mathcal{W} \Gamma E = \langle S, A \rangle \Rightarrow S \Gamma \vdash_{ML} E : A]$$

### 6.2.1 Basic Cases

$$\begin{aligned} \mathcal{W} \Gamma c &= \langle id, B \rangle \\ \text{where } \nu c &= \forall \vec{\varphi}. A \\ B &= A[\vec{\varphi}' / \vec{\varphi}] \\ \text{all } \varphi' &\text{ are fresh} \\ \mathcal{W} \Gamma (\lambda x. E) &= \langle S, S(\varphi \mapsto A) \rangle \\ \text{where } \langle S, A \rangle &= \mathcal{W} (\Gamma, x : \varphi) E \\ \varphi &\text{ is fresh} \end{aligned}$$

$$\begin{aligned} \mathcal{W} \Gamma x &= \langle id, B \rangle \\ \text{where } x : \forall \vec{\varphi}. A &\in \Gamma \\ B &= A[\vec{\varphi}' / \vec{\varphi}] \\ \text{all } \varphi' &\text{ are fresh} \end{aligned}$$

### 6.2.2 Let Construct

$$\begin{aligned} \mathcal{W} \Gamma \text{ (let } x = E_1 \text{ in } E_2) &= \langle S_2 \circ S_1, B \rangle \\ \text{where } \langle S_1, A \rangle &= \mathcal{W} \Gamma E_1 \\ \langle S_2, B \rangle &= \mathcal{W} (S_1 \Gamma, x : \sigma) E_2 \\ \sigma &= S_1 \Gamma A \end{aligned}$$

1. Get the type and substitutions for  $E_1$  given the context  $\Gamma$
2. Get the type and substitutions for  $E_2$ , the context needs to have  $E_1$ 's substitutions applied, we add in a new variable  $x$  (it will be free in  $E_2$ ) and give it a  $\forall$  type that uses no type variables already bound in  $S_1 \Gamma$ .
3. The resulting type for  $E_2$  is the type of the whole term, we must compose the substitutions for  $E_1$  and  $E_2$ .

### 6.2.3 Fix Construct

$$\begin{aligned} \mathcal{W} \Gamma \text{ (fix } g.E) &= \langle S_2 \circ S_1, S_2 A \rangle \\ \text{where } \langle S_1, A \rangle &= \mathcal{W} (\Gamma, g : \varphi) E \\ S_2 &= \text{unify } (S_1 \varphi) A \\ \varphi &\text{ is fresh} \end{aligned}$$

1.  $g$  must have the same type as  $E$  (recursion, the inner call has the same type as the outer), hence to compute the pair for  $E$  we add  $g$  to the context with fresh type variable  $\varphi$
2. We then get a substitution  $S_1$  and type  $A$ , we must unify this with the type of  $g$  (with  $S_1$  applied) to type the whole term.

### 6.2.4 Application

$$\begin{aligned} \mathcal{W} \Gamma (E_1 E_2) &= \langle S_3 \circ S_2 \circ S_1, S_2 \varphi \rangle \\ \text{where } \langle S_1, A \rangle &= \mathcal{W} \Gamma E_1 \\ \langle S_2, B \rangle &= \mathcal{W} (S_1 \Gamma) E_2 \\ S_3 &= \text{unify } (S_2 A) (B \rightarrow \varphi) \\ \varphi &\text{ is fresh} \end{aligned}$$

1. First the type of  $E_1$  is computed as type  $A$ , with substitutions  $S_1$ .
2. Next we get the type of  $E_2$ , first applying the substitution  $S_1$  to the context  $\Gamma$ .
3. We now have  $E_1 : A$  and  $E_2 : B$ .  $A$  must be equal to some type  $B \rightarrow \varphi$  ( $E_1$  is a function taking  $E_2$  as input), hence we unify  $A$  with  $B \rightarrow \varphi$

## 6.3 Polymorphic Recursion

Mycroft generalised Milner's system in an attempt to improve typing for recursively defined objects.

```
map f ls      = if null ls then ls else cons (f (head ls)) (map f (tail ls))
squarelist ls = map (\x -> x^2) ls
```

In  $\Lambda^{NR}$  this would be defined as:

$$\begin{aligned} &\vdots && \text{(definitions of head, tail, cons, etc)} \\ \text{map} &= \lambda f \text{ ls. } \text{Cond } (\text{null ls}) \text{ ls } (\text{cons } (f \text{ (head ls)}) (\text{map } f \text{ (tail ls)})) \\ \text{squarelist} &= \lambda \text{ ls. map } (\lambda x. \text{mul } x \text{ } x) \text{ ls} \end{aligned}$$

The name *squarelist* could then be used in a program.

In ML there is no check to see if functions are independent or mutually recursive, so all definitions must be done in a single step. Hence we can extend  $\mathcal{L}_{ML}$  with a pairing function  $\langle \cdot, \cdot \rangle$ :

let  $\langle \text{map}, \text{squarelist} \rangle = \text{fix } \langle m, s \rangle. \langle \lambda f \text{ ls. } \text{Cond } (\text{null ls}) \text{ l } (\text{cons } (f \text{ (head ls)}) (m \text{ f (tail ls)})), \lambda \text{ ls. } m \text{ } (\lambda x. \text{mul } x \text{ } x) \text{ ls} \rangle$  in ...

However we still have a type assignment issue,  $\mathcal{W}$  will get the following types:

$$\begin{aligned} \text{map} &:: (\text{num} \rightarrow \text{num}) \rightarrow [\text{num}] \rightarrow [\text{num}] \\ \text{squarelist} &:: [\text{num}] \rightarrow [\text{num}] \end{aligned}$$



The definition of map has the type:

$$\text{map} :: \forall \varphi_1 \varphi_2. (\varphi_1 \rightarrow \varphi_2) \rightarrow [\varphi_1] \rightarrow [\varphi_2]$$

For fix  $g.E$  milner's  $\mathcal{W}$  unifies the type of  $E$  and  $g$ , this results in the second type of map not being found by type inference.

One way to avoid this problem is to treat the term as a single definition.

let  $\text{map} = \text{fix } m. \lambda f \text{ ls}. \text{Cond} (\text{null ls}) \text{ ls} (\text{cons } (f (\text{head ls})) (m f (\text{tail ls})))$  in let  $\text{squarelist} = \lambda \text{ls}. \text{map} (\lambda x. \text{mul } x x) \text{ ls}$  in ...

Instead in Mycroft's system the fix rule is altered.

$$\begin{array}{c} \text{(fix)} : \frac{\Gamma, g : A \vdash E : A}{\Gamma \vdash \text{fix } g.E : A} \qquad \text{(fix)} : \frac{\Gamma, g : \tau \vdash_{MYC} E : \tau}{\Gamma \vdash_{MYC} \text{fix } g.E : \tau} \end{array}$$

**Milner's**

**Mycroft's**

Hence the derivation rule allows for type-schemes (the  $\tau$ ) which means different curry types (e.g  $A$ ) may be used.

### 6.3.1 Mycroft-Style Assignment for $\Lambda^{NR}$

The rules for  $\epsilon$ , Call, Rec Call, Def and Rec Def can be replaced by the rules:

$$\begin{array}{c} (\epsilon) : \frac{}{\varepsilon \vdash \epsilon : \diamond} \quad (\text{Call}) : \frac{}{\Gamma; \varepsilon, \text{name} : A \vdash \text{name} : S \ A} \quad (\text{Defs}) : \frac{\varepsilon, \text{name} : A \vdash \text{Defs} : \diamond \quad \emptyset; \varepsilon, \text{name} : A \vdash M : A}{\varepsilon, \text{name} : A \vdash \text{Defs}; \text{name} = M : \diamond} \end{array}$$

With the principle pair algorithm as:

$$\begin{array}{lcl} pp_{\Lambda^{RN}} \ x \ \varepsilon & = & \langle x : \varepsilon; \varepsilon \rangle \text{ where } \varepsilon \text{ is fresh} \\ pp_{\Lambda^{RN}} \ \text{name} \ \varepsilon & = & \langle \emptyset; \text{FreshInstance}(\varepsilon \ \text{name}) \rangle \\ pp_{\Lambda^{RN}} \ (\lambda x.M) \ \varepsilon & = & \begin{cases} \langle \Pi'; A \rightarrow P \rangle & (\Pi = \Pi', x : A) \\ \langle \Pi; \varphi \rightarrow P \rangle & (x \notin \Pi) \end{cases} \\ & & \text{where } \begin{cases} \langle \Pi; P \rangle & = pp_{\Lambda^{RN}} \ M \ \varepsilon \\ \varphi & \text{is fresh} \end{cases} \\ pp_{\Lambda^{RN}} \ (M \ N) \ \varepsilon & = & S_2 \circ S_1 \langle \Pi_1 \cup \Pi_2; \varepsilon \rangle \\ & & \text{where } \begin{cases} \langle \Pi_1; P_1 \rangle & \\ \langle \Pi_2; P_2 \rangle & \end{cases} \quad \begin{array}{l} = pp_{\Lambda^{RN}} \ M \ \varepsilon \\ = pp_{\Lambda^{RN}} \ N \ \varepsilon \end{array} \\ & & \begin{array}{l} S_1 \quad = \text{unify} P_1 \ (P_2 \rightarrow \varphi) \\ S_2 \quad = \text{unifyContexts} (S_1 \ \Pi_1) (S_1 \ \Pi_2) \\ \varphi \quad \text{is fresh} \end{array} \\ pp_{\Lambda^{RN}} \ (\text{Defs}; M) \ \varepsilon & = & \begin{cases} pp_{\Lambda^{RN}} \ M \ \varepsilon & \text{if } \text{CheckEnv} \ \text{Defs} \ \varepsilon \\ \text{untypeable} & \text{otherwise} \end{cases} \\ & & \text{where } \begin{array}{l} \text{CheckEnv} (\text{Defs}; \text{name} = M) \ \varepsilon = (\text{CheckEnv} \ \text{Defs} \ \varepsilon) \wedge (\varepsilon \ \text{name}) = P \\ \text{where } \langle \emptyset, P \rangle = pp_{\Lambda^{RN}} \ M \ \varepsilon \\ \text{CheckEnv} \ \varepsilon \in \quad \quad \quad = \text{true} \end{array} \end{array}$$

### 6.3.2 Milner's and Mycroft's System's Differences

As Mycroft's system is an extension of Milner's:

$$\text{Typeable in Milner's} \Rightarrow \text{Typeable in Mycroft's}$$

- Many terms are typeable in Mycroft's but not in Milner's
- Some terms can be given more general type in Mycroft's than Milner's

The key difference between the systems is that in Milner's recursive calls use the same curry type, but in Mycroft's these can be a more general type, this allows for polymorphic recursion.

Create a term typeable in Mycroft's System but not in Milner's.

$$\text{fix } g.(\lambda(ab.a) (g \lambda c.c) (g \lambda de.d))$$

We can write this as:

```
func :: a -> b
func = (\a b -> a) (func (\c -> c)) (func (\lambda d e -> d))
-- or more idiomatically
func' = const (func' id) (func' const)

-- in execution this looks like:
func x
  = const (func' id) (func' const) x
  = func' id x
  = const (func' id) (func' const) id x
  = func' id id x
  ...
  = func' id id ... id x
```

This is not typeable in Milner's as here  $g$  effectively has two types. Mycroft's allows this as both types can come from the polymorphic type  $\forall \varphi_1 \varphi_2. \varphi_1 \rightarrow \varphi_2$ .

# Chapter 7

## Credit

### Content

Based on the Type Systems course taught by Dr Steffen van Bakel.

These notes were written by Oliver Killane.