We present a formal definition of $\alpha$-equivalence on $\lambda$-terms.

1 Introduction

The terms of the $\lambda$-calculus (or $\lambda$-terms) are given by the following grammar, where $x$ ranges over a countable set $\mathcal{X}$ of variables:

\[
t, u \in \Lambda ::= x \mid \lambda x.t \mid tu
\]

The above grammar leaves implicit the binary application term constructor $@$, in the sense that $tu$ stands for $@\langle t, u \rangle$. In $\lambda x.t$, we say that $\lambda$ is a binder and that $\lambda x$ is the binding site of $x$.

$\lambda$-terms are identified up-to (consistent) renaming of their bound variables. Formally, this identification is made via an equivalence relation on $\Lambda$ called $\alpha$-equivalence and denoted $=_{\alpha}$. For instance we should have

\[
\lambda y. (\lambda x. xy)(xy) =_{\alpha} \lambda y. (\lambda z. zy)(xy)
\]

but not

\[
\lambda y. (\lambda x. xy)(xy) =_{\alpha} \lambda y. (\lambda y. yy)(xy) \quad \text{nor} \quad \lambda y. (\lambda x. xy)(xy) =_{\alpha} \lambda x. (\lambda x. xx)(xx)
\]

Intuitively, $\lambda$-terms quotiented by $\alpha$-equivalence can be represented as directed graphs, in which only free variables are named, while bound variables are represented by edges from their occurrences to their binding site. For instance, the $\alpha$-equivalence class of $\lambda y. (\lambda x. xy)(xy)$ could be represented as the graph depicted in Fig. 1 (left). In §2 we define an algebraic (i.e. term) representation of such graphs.
2 The Locally Nameless Representation

We define $\alpha$-equivalence via an algebraic representation of the graphs such as the one depicted in Fig. 1 (left). This representation uses the well-known De Bruijn indexes to represent bound variables, but still uses names to represent free variables. Such representations are called locally nameless (or sometimes mixed) representations.

The idea of De Bruijn indexes is that an occurrence of a bound variable is represented by a natural number, which counts the number of binders between that occurrence and its binding site. Hence, different occurrences of the same bound variable can be represented by different indexes. For instance, the locally nameless representation of $\lambda y.(\lambda x.x y)(x y)$ is depicted in Fig. 1 (right).

Formally, we consider the following mixed De Bruijn terms, where $n$ ranges over natural numbers:

$$M, N \in \text{\textit{DB}} ::= x \mid n \mid \lambda M \mid MN$$

The mixed De Bruijn representation of a $\lambda$-term $t \in \Lambda$ is the mixed De Bruijn term $t^{\text{db}} \in \text{\textit{DB}}$ defined by induction as follows:

$$
\begin{align*}
    x^{\text{db}} & : = x \\
    (\lambda x.t)^{\text{db}} & : = \lambda (t^{\text{db}}[x \mapsto 0]) \\
    (tu)^{\text{db}} & : = t^{\text{db}} u^{\text{db}}
\end{align*}
$$
where the operation $M[x \mapsto n]$ is defined by induction as

\[
\begin{align*}
    x[x \mapsto n] & := n \\
y[x \mapsto n] & := y \text{ if } y \in \mathcal{X} \text{ and } y \neq x \\
k[x \mapsto n] & := k \text{ if } k \in \mathbb{N} \\
(MN)[x \mapsto n] & := (M[x \mapsto n])(N[x \mapsto n]) \\
(\lambda M)[x \mapsto n] & := \lambda(M[x \mapsto n + 1])
\end{align*}
\]

For instance, we have

\[
(\lambda x.x)^{db} = \lambda((x^{db})[x \mapsto 0])
\]

so that

\[
((\lambda x.x)^{db}) = (\lambda 0)x
\]

and thus

\[
(\lambda x.(\lambda x.x)x)^{db} = \lambda(((\lambda 0)x)[x \mapsto 0]) = \lambda((\lambda 0)0)
\]

Similarly, it is easy to see that in accordance with Fig. 1 (right), we have

\[
(\lambda y.(\lambda x.xy)(xy))^{db} = \lambda((\lambda(01))(x0)) = (\lambda y.(\lambda z.\lambda y)(xy))^{db}
\]

**Definition 2.1** (α-Equivalence). Two λ-terms $t$, $u$ are α-equivalent, notation $t =_\alpha u$, if $t^{db} = u^{db}$.

We refer to [Cha12] for further details and references on the locally nameless representation. For other approaches and references see e.g. [SU06, Chap. 1] or [Kri90, Chap. 1].

**References**

