Co-Debruijn: Everybody's Got To Be Somewhere[2]

From Debruijn to co-Debruijn using Category Theory

Marius Weidner

Chair of Programming Languages, University of Freiburg

January 25, 2024

Outline

1. Getting Started: Scopes and Binders Categorically
   - The Category of Scopes
   - Intrinsically Scoped Debruijn Syntax
2. Going Further: Stronger Intrinsic Invariants over Scopes and Binders
   - The Slice Category of Subscopes
   - A Monad Over Sets Indexed by Scopes
   - Monoidal Structure of Scopes
   - Intrinsically Scopes co-Debruijn Syntax
   - Translating Debruijn Syntax to Codebruijn Syntax
3. Wrapping Up: What I've (Not) Told You
   - This Is Actually an Agda Paper (!)
   - Abstracting over Arbitrary Syntaxes

The Category of Scopes: Δ_+^X

Definition

Let Δ_+^X be the category of scopes.

- Objects: $\bar{x}, \bar{y}, \bar{s} \in |\Delta_+^X| = X^*$
- Morphisms: $f, g \in \Delta_+^X(\bar{x}, \bar{y})$ for $\bar{x}, \bar{y} \in X^*$ are inductively defined:

$$\frac{}{\varepsilon \sqsubseteq \varepsilon}\ \cdot \qquad \frac{\bar{x} \sqsubseteq \bar{y}}{\bar{x}x \sqsubseteq \bar{y}x}\ 1 \qquad \frac{\bar{x} \sqsubseteq \bar{y}}{\bar{x} \sqsubseteq \bar{y}y}\ 0$$

Corollary

The initial object of the Δ_+^X category is the empty scope ε with the $\bar{0}$ as the unique morphism.

Remark

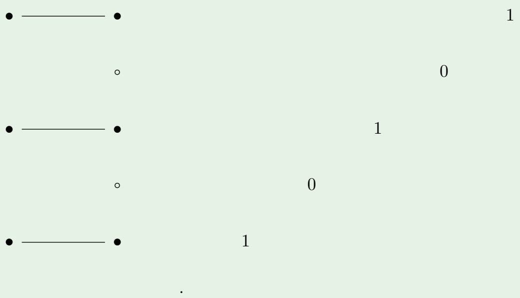
Morphisms in Δ_+^X can be represented by *bit vectors* $\bar{b} \in \{0, 1\}^*$ with one bit per variable of the target scope telling whether it has been mapped to or skipped by the source scope.

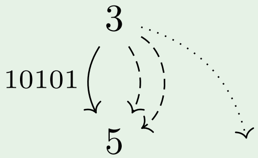
Objects & Morphisms in $\Delta_+^\top$

Example

Let $X = \top$ (where $\top$ is the set with exactly one element $\langle\rangle$).
Thus, Objects $\bar{x} \in X^*$ represents numbers.

$$3 \xrightarrow{10101} 5$$

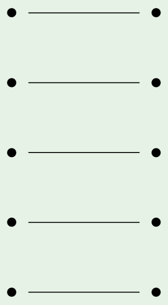




Identity and Composition in $\Delta_+^\top$

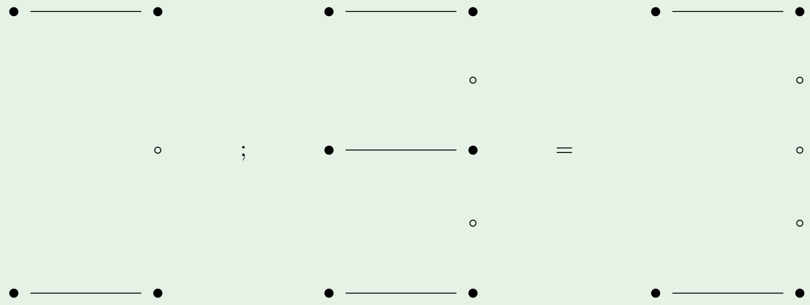
Example

$$5 \xrightarrow{id\ 5} 5$$



Example

$$2 \xrightarrow{101} 3 \quad ; \quad 3 \xrightarrow{10101} 5 \quad = \quad 2 \xrightarrow{10001} 5$$



Δ_+^X is in Fact a Category

Lemma

In Δ_+^X every object $\bar{x} \in X^*$ has an identity morphism, i.e. we can construct an identity morphism for $\bar{x}$ using the inference rules.

Proof.

$$
\begin{aligned}
id \quad & : (\bar{x} : X^*) \to \bar{x} \sqsubseteq \bar{x} \\
id\ \varepsilon & = \cdot \\
id\ \bar{x}x & = (\mathrm{id}\ \bar{x})1
\end{aligned}
$$

□

Corollary

$$
\begin{aligned}
id - l & : id; f = f \\
id - r & : f; id = f
\end{aligned}
$$

Lemma

In Δ_+^X two morphisms $f : \bar{x} \sqsubseteq \bar{y}$ and $g : \bar{y} \sqsubseteq \bar{z}$ compose to a morphism $f;g : \bar{x} \sqsubseteq \bar{z}$, i.e. we can construct a morphism $f;g$ from f and g using the inference rules.

Proof.

$$
\begin{aligned}
\_;\_ \quad & : \bar{x} \sqsubseteq \bar{y} \to \bar{y} \sqsubseteq \bar{z} \to \bar{x} \sqsubseteq \bar{z} \\
\cdot \ ; \ \cdot & = \cdot \\
f1 \ ; \ g1 & = (f;g)1 \\
f0 \ ; \ g1 & = (f;g)0 \\
f \ ; \ g0 & = (f;g)0
\end{aligned}
$$

□

Corollary

$$
\begin{aligned}
\textit{assoc} \quad & : f;(g;h) = (f;g);h \\
\textit{antisym} & : (f : \bar{x} \sqsubseteq \bar{y}) \to (g : \bar{y} \sqsubseteq \bar{z}) \to \bar{x} = \bar{y} \wedge f = g = id\ \bar{x}
\end{aligned}
$$

Intrinsically Scoped Debruijn Syntax via $\Delta_{+}^{\top}$

Definition

Let $Tm : |\Delta_{+}^{\top}| \rightarrow Set$ be inductively defined:

$$\frac{\langle\rangle \sqsubseteq \bar{x}}{Tm\ \bar{x}}\ \# \qquad \frac{Tm\ \bar{x} \quad Tm\ \bar{x}}{Tm\ \bar{x}}\ \$ \qquad \frac{Tm\ \bar{x}\langle\rangle}{Tm\ \bar{x}}\ \lambda$$

Example

Lifting Scope Indexed Terms using Composition in $\Delta_+^\top$

Lemma

Given an intrinsically scoped term $t \in Tm\ \bar{x}$ we can lift t to a $Tm\ \bar{y}$, if there exists a morphism $\bar{x} \sqsubseteq \bar{y} \in \Delta_+^\top(\bar{x}, \bar{y})$, i.e. $\bar{x}$ is a subscope of $\bar{y}$.

Proof.

$$
\begin{array}{lll}
\_ \uparrow \_ & : & Tm\ \bar{x} \to \bar{x} \sqsubseteq \bar{y} \to Tm\ \bar{y} \\
(\#\ v) \quad \uparrow f & = & \#\ (v; f) \\
(t_1\ \$\ t_2) \uparrow f & = & (t_1 \uparrow f)\ \$\ (t_2 \uparrow f) \\
(\lambda\ t) \quad \uparrow f & = & \lambda\ (t \uparrow S f)
\end{array}
$$

□

The Slice Category of Subscopes: $\Delta_+^X \setminus \bar{s}$

Definition

Let $\Delta_+^X \setminus \bar{s}$ be the category of subscopes for a given $\bar{s} \in X^*$.

- Objects: $\bar{b}, (\bar{x}, f) \in |\Delta_+^X \setminus \bar{s}| = (\bar{x} : X^* \times \Delta_+^X(\bar{x}, \bar{s}))$
- Morphisms: $h \in [\Delta_+^X \setminus \bar{s}]((\bar{x}, f), (\bar{y}, g))$ such that $f = h; g$

$$
\begin{tikzcd}
\bar{x} \arrow[rr, "h"] \arrow[dr, dotted, "f"'] & & \bar{y} \arrow[dl, dotted, "g"] \\
& \bar{s} &
\end{tikzcd}
$$

Corollary

The initial object of the $\Delta_+^X \setminus \bar{s}$ category is the empty subscope $(\varepsilon, \bar{0})$.

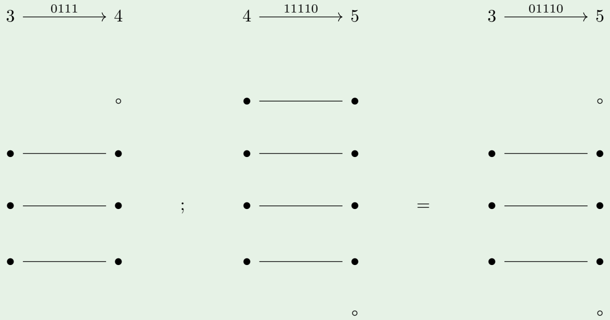
Remark

Objects in $\Delta_+^X \setminus \bar{s}$ can be represented by *bit vectors* $\bar{b} \in \{0, 1\}^*$ with one bit per variable of scope $\bar{s}$, telling whether it has been selected.

Objects & Morphisms in $\Delta_+^T \setminus 5$

Example

$$
\begin{tikzcd}
3 \arrow[rr, "0111"] \arrow[dr, dotted, "01110"'] & & 4 \arrow[dl, dotted, "11110"] \\
& 5 &
\end{tikzcd}
$$



Alternatively:

- $(3, 01110) \xrightarrow{0111} (4, 11110)$
- $01110 \xrightarrow{0111} 11110$

Coproducts in $\Delta_+^X \setminus \bar{s}$

Theorem

Objects in the slice category $\bar{b}_1, \bar{b}_2 \in |\Delta_+^X \setminus \bar{s}|$ *have a coproduct object* $\bar{b}_1 + \bar{b}_2$ *if there exist morphisms* $l \in [\Delta_+^X \setminus \bar{s}](\bar{b}_1, \bar{b}_1 + \bar{b}_2)$ *and* $r \in [\Delta_+^X \setminus \bar{s}](\bar{b}_2, \bar{b}_1 + \bar{b}_2)$*. Then for every* $f \in [\Delta_+^X \setminus \bar{s}](\bar{b}_1, \bar{b}_3)$ *and* $g \in [\Delta_+^X \setminus \bar{s}](\bar{b}_2, \bar{b}_3)$*, there exists a unqiue* $h \in [\Delta_+^X \setminus \bar{s}](\bar{b}_1 + \bar{b}_2, \bar{b}_3)$ *such that* $f = l; h$ *and* $g = r; h$*.*

Example

$$\begin{tikzcd}
01100 \arrow[dr, "011"'] \arrow[drr, "0110"] & & \\
& 11100 \arrow[r, "1110"] & 11110 \\
11000 \arrow[ur, "110"] \arrow[urr, "1100"'] & &
\end{tikzcd}$$

Remark

The coproduct $\bar{b}_1 + \bar{b}_2$ of two subscopes $\bar{b}_1, \bar{b}_2$ corresponds to the minimal subscope covering both $\bar{b}_1$ and $\bar{b}_2$. The coproduct $\bar{b}_1 + \bar{b}_2$ can be computed by pointswise disjunction of $\bar{b}_1$ and $\bar{b}_2$.

Category of Sets Indexed by Scopes

Definition

Let Set_X be the category of sets indexed by scopes $\bar{x} \in X^*$.

- Objects: $T, S \in |Set_X| = X^* \to Set = \bar{X}$
- Morphisms: $f_{\cdot} \in Set_X(T, S) = (\bar{x} \in X^*) \to T(\bar{x}) \to S(\bar{x}) = T \dot{\to} S$

Definition

Let $\_ \Uparrow \_ : \bar{X} \to \bar{X} = (T, \bar{x}) \mapsto (T(\bar{s}) \times \bar{s} \sqsubseteq \bar{x})$.

We define $Ref : Set_X \dot{\to} Set_X$ to be the endofunctor induced by the mapping

- $Ref(T) = \bar{x} \mapsto T \Uparrow \bar{x} \in \bar{X}$
- $Ref(f_{\cdot}) = (t, h) \mapsto (f_{\cdot}(t), h) \in T \dot{\to} S$

Remark

The set $T \Uparrow \bar{x}$ packs an set $T \in \bar{X}$ indexed by $\bar{x} \in X^*$ applied to a subscope $\bar{s}$ of $\bar{x}$, together with a selection $h \in |\Delta_+^X \setminus \bar{x}|$ of the variables of T.

Δ_{+}^{X} makes Ref a Monad!

Theorem

The functor $Ref : Set_X \rightarrow Set_X$ gives rise to a monad with the two natural transformations

- $unit : Id(T) \stackrel{.}{\rightarrow} Ref(T) = t \mapsto (t, id)$
- $mult : Ref(Ref(T)) \stackrel{.}{\rightarrow} Ref(T) = ((t, h_1)h_2) \mapsto (t, h_1 ; h_2)$

Example

$$
\begin{tikzcd}
 & Tm \arrow[r, dashed, "unit"] & \bar{x} \mapsto (Tm\ \bar{x}, \bar{x} \sqsubseteq \bar{x}) \arrow[dll, dashed, "Ref"'] \\
\bar{y} \mapsto ([\bar{x} \mapsto (Tm\ \bar{x}, \bar{x} \sqsubseteq \bar{x})]\bar{s}, \bar{s} \sqsubseteq \bar{y}) \arrow[rr, dashed, "mult"] & & \bar{y} \mapsto (Tm\ \bar{s}, \bar{s} \sqsubseteq \bar{y})
\end{tikzcd}
$$

The Notion of Relevant Pairs

Definition

Let $\_ \times_R \_ : \bar{X} \to \bar{X} \to \bar{X} = (T, S, \bar{x}) \mapsto ((\_, \overline{b}_1 : T \Uparrow \bar{x}) \times (\_, \overline{b}_2 : S \Uparrow \bar{x}) \times \exists \overline{b}(\overline{b} \equiv \overline{b}_1 + \overline{b}_2))$.
with $\_ ,_R \_ : T \Uparrow \bar{x} \to S \Uparrow \bar{x} \to (T \times_R S) \Uparrow \bar{x} = ((t_1, \overline{b}_1), (t_2, \overline{b}_2)) \mapsto ((t_1, \overline{b}'_1), (t_2, \overline{b}'_2), (\overline{b}_1 | \overline{b}_2))$
to construct a $T \times_R S$, where $\overline{b}'_{1,2} = ((\overline{b}_1 | \overline{b}_2) \& \overline{b}_{1,2}) \upharpoonright \Sigma(\overline{b}_1 | \overline{b}_2)$

Example

Let $(t_1, \overline{b}_1) : Tm\ 2 \times 2 \sqsubseteq 4 = .., 01100)$
and $(t_2, \overline{b}_2) : Tm\ 2 \times 2 \sqsubseteq 4 = .., 11000)$.
Then $\overline{b} : 3 \sqsubseteq 4 = \overline{b}_1 | \overline{b}_2 = 11100$. and $\overline{b}'_1 =$

Example

$$\begin{tikzcd}
01100 \arrow[dr, "011"'] \arrow[drr, "0110"] & & \\
& 11100 \arrow[r, "1110"] & 11110 \\
11000 \arrow[ur, "110"] \arrow[urr, "1100"'] & &
\end{tikzcd}$$

References

[1] Conor McBride. *Cats and types: Best friends?* Aug. 2021. URL: https://www.youtube.com/watch?v=05IJ3YL8p0s.

[2] Conor McBride. "Everybody's Got To Be Somewhere". In: *Electronic Proceedings in Theoretical Computer Science* 275 (July 2018), pp. 53–69. ISSN: 2075-2180. DOI: 10.4204/eptcs.275.6. URL: http://dx.doi.org/10.4204/EPTCS.275.6.