A Solution to the POPLMARK Challenge in Isabelle/HOL

Stefan Berghofer

August 16, 2018

Abstract

We present a solution to the POPLMARK challenge designed by Aydemir et al., which has as a goal the formalization of the meta-theory of System $F_{<}$. The formalization is carried out in the theorem prover Isabelle/HOL using an encoding based on de Bruijn indices. We start with a relatively simple formalization covering only the basic features of System $F_{<}$, and explain how it can be extended to also cover records and more advanced binding constructs.

Contents

1 General Utilities 2

2 Formalization of the basic calculus 3
  2.1 Types and Terms .............................................. 3
  2.2 Lifting and Substitution ..................................... 4
  2.3 Well-formedness .............................................. 7
  2.4 Subtyping .................................................. 10
  2.5 Typing ......................................................... 12
  2.6 Evaluation .................................................. 14

3 Extending the calculus with records 15
  3.1 Types and Terms .............................................. 15
  3.2 Lifting and Substitution ..................................... 17
  3.3 Well-formedness .............................................. 25
  3.4 Subtyping .................................................. 27
  3.5 Typing ......................................................... 28
  3.6 Evaluation .................................................. 32

4 Evaluation contexts 36

5 Executing the specification 38
1 General Utilities

This section introduces some general utilities that will be useful later on in the formalization of System F<. The following rewrite rules are useful for simplifying mutual induction rules.

**Lemma True-simps:**
\[
(\text{True} \implies \text{PROP } P) \equiv \text{PROP } P \\
(\text{PROP } P \implies \text{True}) \equiv \text{PROP } \text{Trueprop True} \\
(\forall x. \text{True}) \equiv \text{PROP } \text{Trueprop True}
\]

(\text{proof})

Unfortunately, the standard introduction and elimination rules for bounded universal and existential quantifier do not work properly for sets of pairs.

**Lemma ballpI:**
\[
(\forall x, y. (x, y) \in A \implies P x y) \implies \forall (x, y) \in A. P x y
\]

(\text{proof})

**Lemma bspec:**
\[
\forall (x, y) \in A. P x y \implies (x, y) \in A \implies P x y
\]

(\text{proof})

**Lemma ballpE:**
\[
\forall (x, y) \in A. P x y \implies (P x y \implies Q) \implies \\
((x, y) \notin A \implies Q) \implies Q
\]

(\text{proof})

**Lemma bexpI:**
\[
P x y \implies (x, y) \in A \implies \exists (x, y) \in A. P x y
\]

(\text{proof})

**Lemma bexpE:**
\[
\exists (x, y) \in A. P x y \implies \\
(\forall x, y. (x, y) \in A \implies P x y \implies Q) \implies Q
\]

(\text{proof})

**Lemma ball-eq-sym:**
\[
\forall (x, y) \in S. f x y = g x y \implies \forall (x, y) \in S. g x y = f x y
\]

(\text{proof})

**Lemma wf-measure-size:**
\[
\text{wf} (\text{measure size})
\]

(\text{proof})

**Notation**

\begin{itemize}
\item Some (\lfloor \cdot \rfloor)
\item None (\perp)
\item length (\|\cdot\|)
\item Cons (\cdot ::/ \cdot [66, 65] 65)
\end{itemize}

The following variant of the standard nth function returns \perp if the index is
out of range.

**primrec**

\[
th{el} :: 'a list ⇒ nat ⇒ 'a option (\(-\)) [90, 0] 91
\]

**where**

\[[(i) = \bot | \ \{ x \# xs\}(i) = (case i of 0 ⇒ [x] | Suc j ⇒ xs \(j\))]\]

**lemma [simp]:** \[i < \|xs\| ⇒ (xs @ ys)(i) = xs\(i)\]

**lemma [simp]:** \[\|xs\| ≤ i ⇒ (xs @ ys)(i) = ys\(i - \|xs\|)\]

**Association lists**

**primrec** assoc :: ('a × 'b) list ⇒ 'a ⇒ 'b option (\(-\)) [90, 0] 91

**where**

\[[(a)? = \bot | \ \{ x \# xs\}(a)? = (if fst x = a then \[snd x\] else xs\(a)?)\]

**primrec** unique :: ('a × 'b) list ⇒ bool

**where**

\[unique [] = True | \ unique (x # xs) = (xs\(fst x)? = \bot ∧ unique xs)\]

**lemma** assoc-set: \[ps(x)? = \[y\] ⇒ (x, y) ∈ set ps\]

**lemma** map-assoc-None [simp]: \[ps(x)? = \bot ⇒ map (\(x, y\). (x, f x y)) ps\(x)? = \bot\]

**no-syntax**

- Map :: maplets ⇒ 'a → 'b (([[-]])

## 2 Formalization of the basic calculus

In this section, we describe the formalization of the basic calculus without records. As a main result, we prove type safety, presented as two separate theorems, namely preservation and progress.

### 2.1 Types and Terms

The types of System F\(\langle\rangle\) are represented by the following datatype:

**datatype type =**
The subtyping and typing judgements depend on a context (or environment) \( \Gamma \) containing bindings for term and type variables. A context is a list of bindings, where the \( i \)th element \( \Gamma[i] \) corresponds to the variable with index \( i \).

The following datatype represents the terms of System \( \text{F}_{<:} \):

\[
\text{trm} = \begin{align*}
& \text{Var} \ \text{nat} \\
| & \text{Abs} \ \text{type} \ \text{trm} \ ((\lambda x : \cdot \cdot) [0, 10] 10) \\
| & \text{TAbs} \ \text{type} \ \text{trm} \ ((\lambda<: \cdot \cdot) [0, 10] 10) \\
| & \text{App} \ \text{trm} \ \text{trm} \ (\text{infixl} \cdot 200) \\
| & \text{TApp} \ \text{trm} \ \text{type} \ (\text{infixl} \cdot, 200)
\end{align*}
\]

### 2.2 Lifting and Substitution

One of the central operations of \( \lambda \)-calculus is substitution. In order to avoid that free variables in a term or type get “captured” when substituting it for a variable occurring in the scope of a binder, we have to increment the indices of its free variables during substitution. This is done by the lifting
functions $\uparrow_T n k$ and $\uparrow n k$ for types and terms, respectively, which increment the indices of all free variables with indices $\geq k$ by $n$. The lifting functions on types and terms are defined by

\begin{verbatim}
promrec liftT :: nat ⇒ nat ⇒ type ⇒ type (↑_T)
  where
  ↑_T n k (TVar i) = (if i < k then TVar i else TVar (i + n))
  ↑_T n k Top = Top
  ↑_T n k (T ⇒ U) = ↑_T n k T ⇒ ↑_T n k U
  ↑_T n k (∀T. U) = (∀↑_T n k T. ↑_T n (k + 1) U)

promrec lift :: nat ⇒ nat ⇒ trm ⇒ trm (↑)
  where
  ↑ n k (Var i) = (if i < k then Var i else Var (i + n))
  ↑ n k (λT. t) = (λ↑ n k T. ↑ n (k + 1) t)
  ↑ n k (s · t) = ↑ n k s · ↑ n k t
  ↑ n k (t · r) = ↑ n k T
\end{verbatim}

It is useful to also define an “unlifting” function $\downarrow_T n k$ for decrementing all free variables with indices $\geq k$ by $n$. Moreover, we need several substitution functions, denoted by $T[k הו_Τ S]_T$, $t[k הו_Τ S]_T$, and $t[k הו s]$, which substitute type variables in types, type variables in terms, and term variables in terms, respectively. They are defined as follows:

\begin{verbatim}
promrec substTT :: type ⇒ nat ⇒ type ⇒ type (\ [-] T_T_T
  where
  (TVar i)[k ↦_Τ S]_Τ = (if i < k then TVar (i - 1) else if i = k then ↑_T k 0 S else TVar i)
  Top[k ↦_Τ S]_Τ = Top
  (T ⇒ U)[k ↦_Τ S]_Τ = T[k ↦_Τ S]_Τ ⇒ U[k ↦_Τ S]_Τ
  (∀T. U)[k ↦_Τ S]_Τ = (∀T[k ↦_Τ S]_Τ. U[k + 1 ↦_Τ S]_Τ)

promrec decT :: nat ⇒ nat ⇒ type ⇒ type (\_ T)
  where
  ↓_T 0 k T = T
  ↓_T (Suc n) k T = ↓_T n k (T[k ↦_Τ Top]_Τ)

promrec subst :: trm ⇒ nat ⇒ trm ⇒ trm (\ [-] T_T)
  where
  (Var i)[k ↦ s] = (if i < k then Var (i - 1) else if i = k then ↑ 0 s else Var i)
  (t · u)[k ↦ s] = t[k ↦ s] · u[k ↦ s]
  (λT. t)[k ↦ s] = (λ↑ k T. t[k + 1 ↦ s])
  (λ<:T. t)[k ↦ s] = (λ<: k T. t[k + 1 ↦ s])

promrec substT :: trm ⇒ nat ⇒ type ⇒ trm (\ [-] T_T
  where
  (Var i)[k ↦ S] = (if i < k then Var (i - 1) else Var i)
  (t · u)[k ↦ S] = t[k ↦ S] · u[k ↦ S]
\end{verbatim}

5
Lifting and substitution extends to typing contexts as follows:

\[
\begin{align*}
\text{primrec } & \text{lifte} :: \text{nat } \Rightarrow \text{nat } \Rightarrow \text{env } \Rightarrow \text{env } \\
\text{where} & \quad \uparrow e n k \[] = \[] \\
\text{primrec } & \text{subst} :: \text{env } \Rightarrow \text{nat } \Rightarrow \text{type } \Rightarrow \text{env } \\
\text{where} & \quad [\[ k \mapsto \tau T \] e] = \[] \\
\text{primrec } & \text{dec} :: \text{nat } \Rightarrow \text{nat } \Rightarrow \text{env } \Rightarrow \text{env } \\
\text{where} & \quad \downarrow 0 k \Gamma = \Gamma \\
\text{where} & \quad \downarrow (\text{Suc } n) k \Gamma = \downarrow n k (\Gamma[k \mapsto \text{Top}])
\end{align*}
\]

Note that in a context of the form \(B :: \Gamma\), all variables in \(B\) with indices smaller than the length of \(\Gamma\) refer to entries in \(\Gamma\) and therefore must not be affected by substitution and lifting. This is the reason why an additional offset \(\|\Gamma\|\) needs to be added to the index \(k\) in the second clauses of the above functions. Some standard properties of lifting and substitution, which can be proved by structural induction on terms and types, are proved below. Properties of this kind are quite standard for encodings using de Bruijn indices and can also be found in papers by Barras and Werner [2] and Nipkow [3].

**Lemma** lifte-length [simp]: \(\|\uparrow e n k\| = \|\Gamma\|\)

**Lemma** substE-length [simp]: \(\|\Gamma[k \mapsto U] e\| = \|\Gamma\|\)

**Lemma** lifte-nth [simp]:

\[
\begin{align*}
\uparrow e n k \Gamma(i) &= \text{map-option} (\uparrow n (k + \|\Gamma\| - i - 1)) \Gamma(i) \\
\downarrow e k \Gamma(i) &= \text{map-option} (\lambda U. U[\|\Gamma\| - i \mapsto T]_e) \Gamma(i)
\end{align*}
\]

**Lemma** substE-nth [simp]:

\[
\begin{align*}
\Gamma[0 \mapsto T]_e(i) &= \text{map-option} (\lambda U. U[\|\Gamma\| - i - 1 \mapsto T]_e) \Gamma(i) \\
\Gamma[0 \mapsto T']_e(i) &= \text{map-option} (\lambda U. U[\|\Gamma\| - i - 1 \mapsto T']_e) \Gamma(i)
\end{align*}
\]

**Lemma** liftT-liftT [simp]:

\[
i \leq j \Rightarrow j \leq i + m \Rightarrow \uparrow n j (\uparrow m i T) = \uparrow (m + n) i T
\]

**Lemma** liftT-liftT' [simp]:

\[
6
\]
\[i + m \leq j \implies \uparrow n j (\uparrow m i T) = \uparrow m i (\uparrow n (j - m) T)\]

\textbf{lemma} lift-size \[\text{simp}]:\ size (\uparrow n k T) = size T

\textbf{lemma} liftT0 \[\text{simp}]: \uparrow 0 i T = T

\textbf{lemma} lift0 \[\text{simp}]: \uparrow 0 i t = t

\textbf{theorem} substT-liftT \[\text{simp}]:
\[
k \leq k' \implies k' < k + n \implies (\uparrow n k T)[k' \mapsto \tau U] = \uparrow (n - 1) k T
\]

\textbf{theorem} liftT-substT \[\text{simp}]:
\[
k \leq k' \implies \uparrow n k (T[k' \mapsto \tau U]) = \uparrow n k T[k' + n \mapsto \tau U]
\]

\textbf{theorem} liftT-substT' \[\text{simp}]:
\[
k' < k \implies \uparrow n k (T[k' \mapsto \tau U]) = \uparrow n (k + 1) T[k' \mapsto \tau \uparrow n (k - k') U]
\]

\textbf{lemma} liftT-substT-Top \[\text{simp}]:
\[
k \leq k' \implies \uparrow n k' (T[k \mapsto \tau \text{Top}]) = \uparrow n (\text{Suc} k') T[k \mapsto \tau \text{Top}]
\]

\textbf{lemma} lift-lift \[\text{simp}]:
\[
k \leq k' \implies k' \leq k + n \implies \uparrow n' k' (\uparrow n k t) = \uparrow (n + n') k t
\]

\textbf{lemma} substT-substT:
\[
i \leq j \implies T[Suc j \mapsto \tau V][i \mapsto \tau U][j - i \mapsto \tau V] = T[i \mapsto \tau U][j \mapsto \tau V]
\]

\section{2.3 Well-formedness}

The subtyping and typing judgements to be defined in §2.4 and §2.5 may only operate on types and contexts that are well-formed. Intuitively, a type \(T\) is well-formed with respect to a context \(\Gamma\), if all variables occurring in it are defined in \(\Gamma\). More precisely, if \(T\) contains a type variable \(\text{TVar} i\), then the \(i\)th element of \(\Gamma\) must exist and have the form \(\text{TVar} B U\).

\textit{inductive}
**well-formed :: env ⇒ type ⇒ bool** (-wf - [50, 50] 50)

**where**

- **wf-TVar**: \( \Gamma(i) = [\text{TVar}\ B\ T] \implies \Gamma \vdash_{\text{wf}} \text{TVar}\ i \)
- **wf-Top**: \( \Gamma \vdash_{\text{wf}} \text{Top} \)
- **wf-arrow**: \( \Gamma \vdash_{\text{wf}} \ T \implies \Gamma \vdash_{\text{wf}} U \implies \Gamma \vdash_{\text{wf}} T \rightarrow U \)
- **wf-all**: \( \Gamma \vdash_{\text{wf}} \ T \implies \text{TVar}\ B\ T :: \Gamma \vdash_{\text{wf}} U \implies \Gamma \vdash_{\text{wf}} (\forall \ <\ : T.\ U) \)

A context \( \Gamma \) is well-formed, if all types occurring in it only refer to type variables declared “further to the right”:

**inductive**

**well-formedE :: env ⇒ bool** (-wf [50] 50)

**and** well-formedB :: env ⇒ binding ⇒ bool (-wfB - [50, 50] 50)

**where**

\[ \Gamma \vdash_{\text{wfB}} B \equiv \Gamma \vdash_{\text{wf}} \text{type-ofB}\ B \]

**wf-Nil**: \([]\) \( \vdash_{\text{wf}} \)

**wf-Cons**: \( \Gamma \vdash_{\text{wfB}} B \implies \Gamma \vdash_{\text{wf}} B :: \Gamma \vdash_{\text{wf}} \)

The judgement \( \Gamma \vdash_{\text{wfB}} B \), which denotes well-formedness of the binding \( B \) with respect to context \( \Gamma \), is just an abbreviation for \( \Gamma \vdash_{\text{wf}} \text{type-ofB}\ B \). We now present a number of properties of the well-formedness judgements that will be used in the proofs in the following sections.

**inductive-cases well-formed-cases**:

- \( \Gamma \vdash_{\text{wf}} \text{TVar}\ i \)
- \( \Gamma \vdash_{\text{wf}} \text{Top} \)
- \( \Gamma \vdash_{\text{wf}} T \rightarrow U \)
- \( \Gamma \vdash_{\text{wf}} (\forall \ <\ : T.\ U) \)

**inductive-cases well-formedE-cases**:

- \( B :: \Gamma \vdash_{\text{wf}} \)

**lemma** \( \text{wf-TVarB}: \Gamma \vdash_{\text{wf}} T \implies \Gamma \vdash_{\text{wf}} \text{TVar}\ B\ T :: \Gamma \vdash_{\text{wf}} \)  
  \( \text{proof} \)

**lemma** \( \text{wf-VarB}: \Gamma \vdash_{\text{wf}} T \implies \Gamma \vdash_{\text{wf}} \text{Var}\ B\ T :: \Gamma \vdash_{\text{wf}} \)  
  \( \text{proof} \)

**lemma** \( \text{map-is-TVarb}: \text{map is-TVarB}\ \Gamma' = \text{map is-TVarB}\ \Gamma \implies \Gamma(i) = [\text{TVar}\ B\ T] \implies \exists T.\ \Gamma'(i) = [\text{TVar}\ B\ T] \)  
  \( \text{proof} \)

A type that is well-formed in a context \( \Gamma \) is also well-formed in another context \( \Gamma' \) that contains type variable bindings at the same positions as \( \Gamma \):

**lemma** \( \text{wf-equal-length}: \)  
  \( \text{assumes } H: \Gamma \vdash_{\text{wf}} T \)
  \( \text{shows } \text{map is-TVarB}\ \Gamma' = \text{map is-TVarB}\ \Gamma \implies \Gamma' \vdash_{\text{wf}} T \)  
  \( \text{proof} \)

A well-formed context of the form \( \Delta \ @\ B :: \Gamma \) remains well-formed if we replace the binding \( B \) by another well-formed binding \( B' \):
**lemma** $\text{wfE-replace}$:

$$
\begin{align*}
\Delta @ B :: \Gamma \vdash_{wf} & \Gamma \vdash_{wf} B'B = \text{is-Var} B B' \\
\Delta @ B' :: \Gamma \vdash_{wf} & 
\end{align*}
$$

The following weakening lemmas can easily be proved by structural induction on types and contexts:

**lemma** $\text{wf-weaken}$:

assumes $H: \Delta @ \Gamma \vdash_{wf} T$

shows $\uparrow_e (\text{Suc 0}) 0 \Delta @ B :: \Gamma \vdash_{wf} \uparrow_\tau (\text{Suc 0}) \|\Delta\| T$

(proof)

**lemma** $\text{wf-weaken'}$: $\Gamma \vdash_{wf} T \Rightarrow \Delta @ \Gamma \vdash_{wf} \uparrow_\tau \|\Delta\| 0 T$

(proof)

**lemma** $\text{wfE-weaken}$: $\Delta @ \Gamma \vdash_{wf} \Rightarrow \Gamma \vdash_{wf} B \Rightarrow \uparrow_e (\text{Suc 0}) 0 \Delta @ B :: \Gamma \vdash_{wf} B$

(proof)

Intuitively, lemma $\text{wf-weaken}$ states that a type $T$ which is well-formed in a context is still well-formed in a larger context, whereas lemma $\text{wfE-weaken}$ states that a well-formed context remains well-formed when extended with a well-formed binding. Owing to the encoding of variables using de Bruijn indices, the statements of the above lemmas involve additional lifting functions. The typing judgement, which will be described in §2.5, involves the lookup of variables in a context. It has already been pointed out earlier that each entry in a context may only depend on types declared “further to the right”. To ensure that a type $T$ stored at position $i$ in an environment $\Gamma$ is valid in the full environment, as opposed to the smaller environment consisting only of the entries in $\Gamma$ at positions greater than $i$, we need to increment the indices of all free type variables in $T$ by $\text{Suc } i$:

**lemma** $\text{wf-liftB}$:

assumes $H: \Gamma \vdash_{wf} T$

shows $\Gamma(i) = [\text{Var} B T] \Rightarrow \Gamma \vdash_{wf} \uparrow_\tau (\text{Suc 0}) 0 T$

(proof)

We also need lemmas stating that substitution of well-formed types preserves the well-formedness of types and contexts:

**theorem** $\text{wf-subst}$:

$$
\begin{align*}
\Delta @ B :: \Gamma \vdash_{wf} U \Rightarrow \Delta[\theta \mapsto U] \varepsilon @ \Gamma \vdash_{wf} T[\|\Delta\| \mapsto_\tau U]_\tau
\end{align*}
$$

(proof)

**theorem** $\text{wfE-subst}$: $\Delta @ B :: \Gamma \vdash_{wf} U \Rightarrow \Delta[\theta \mapsto U]_\varepsilon @ \Gamma \vdash_{wf}$

(proof)
2.4 Subtyping

We now come to the definition of the subtyping judgement \( \Gamma \vdash T <: U \).

**inductive**

\[
\text{subtyping} :: \text{env} \Rightarrow \text{type} \Rightarrow \text{type} \Rightarrow \text{bool}
\]

\[
(\vdash - <: - [50, 50, 50] 50)
\]

**where**

\[
\begin{align*}
\text{SA-Top}: & \quad \Gamma \vdash wf \implies \Gamma \vdash w_f S \implies \Gamma \vdash S <: Top \\
| \text{SA-refl-TVar}: & \quad \Gamma \vdash w_f \implies \Gamma \vdash w_f TVar i \implies \Gamma \vdash TVar i <: TVar i \\
| \text{SA-trans-TVar}: & \quad \Gamma \vdash \tau \langle i \rangle = \lfloor TVarB U \rfloor \implies \\
& \quad \Gamma \vdash \tau \langle \text{Suc} i \rangle 0 U <: T \implies \Gamma \vdash TVar i <: T \\
| \text{SA-arrow}: & \quad \Gamma \vdash T_1 <: S_1 \implies \Gamma \vdash S_2 <: T_2 \implies \Gamma \vdash S_1 \rightarrow S_2 <: T_1 \rightarrow T_2 \\
| \text{SA-all}: & \quad \Gamma \vdash (\forall <: S_1. S_2) <: (\forall <: T_1. T_2)
\end{align*}
\]

The rules \text{SA-Top} and \text{SA-refl-TVar}, which appear at the leaves of the derivation tree for a judgement \( \Gamma \vdash T <: U \), contain additional side conditions ensuring the well-formedness of the contexts and types involved. In order for the rule \text{SA-trans-TVar} to be applicable, the context \( \Gamma \) must be of the form \( \Gamma_1 @ B :: \Gamma_2 \), where \( \Gamma_1 \) has the length \( i \). Since the indices of variables in \( B \) can only refer to variables defined in \( \Gamma_2 \), they have to be incremented by \( \text{Suc} i \) to ensure that they point to the right variables in the larger context \( \Gamma \).

**lemma** \text{wf-subtype-env}:

\[
\begin{align*}
& \text{assumes} \ P Q: \Gamma \vdash P <: Q \\
& \text{shows} \ \Gamma \vdash w_f (\text{proof})
\end{align*}
\]

**lemma** \text{wf-subtype}:

\[
\begin{align*}
& \text{assumes} \ P Q: \Gamma \vdash P <: Q \\
& \text{shows} \ \Gamma \vdash w_f P \land \Gamma \vdash w_f Q (\text{proof})
\end{align*}
\]

**lemma** \text{wf-subtypeE}:

\[
\begin{align*}
& \text{assumes} \ H: \Gamma \vdash T <: U \\
& \text{and} \ H \vdash \Gamma \vdash w_f \implies \Gamma \vdash w_f T \implies \Gamma \vdash w_f U \implies P \\
& \text{shows} \ P (\text{proof})
\end{align*}
\]

By induction on the derivation of \( \Gamma \vdash T <: U \), it can easily be shown that all types and contexts occurring in a subtyping judgement must be well-formed:

**lemma** \text{wf-subtype-conj}:

\[
\Gamma \vdash T <: U \implies \Gamma \vdash w_f \land \Gamma \vdash w_f T \land \Gamma \vdash w_f U (\text{proof})
\]

By induction on types, we can prove that the subtyping relation is reflexive:

**lemma** \text{subtype-refl}:

\[
\Gamma \vdash w_f \implies \Gamma \vdash w_f T \implies \Gamma \vdash T <: T (\text{proof})
\]
The weakening lemma for the subtyping relation is proved in two steps: by induction on the derivation of the subtyping relation, we first prove that inserting a single type into the context preserves subtyping:

**Lemma subtype-weaken:**

assumes \( \Delta @ \Gamma \vdash P <: Q \)

and \( \Gamma \vdash_{w_f} B \)

shows \( \uparrow_e 1 \ 0 \ \Delta @ B :: \Gamma \vdash \uparrow_\tau 1 \ \| \Delta \| \ 0 P <: \uparrow_\tau 1 \ \| \Delta \| \ 0 Q \)

All cases are trivial, except for the \( SA-trans-TVar \) case, which requires a case distinction on whether the index of the variable is smaller than \( \| \Delta \| \).

The stronger result that appending a new context \( \Delta \) to a context \( \Gamma \) preserves subtyping can be proved by induction on \( \Delta \), using the previous result in the induction step:

**Lemma subtype-weaken': — A.2**

\( \Gamma \vdash P <: Q \Rightarrow \Delta @ \Gamma \vdash_{w_f} \Rightarrow \Delta @ \Gamma \vdash \uparrow_\tau \ 0 P <: \uparrow_\tau \ 0 Q \)

An unrestricted transitivity rule has the disadvantage that it can be applied in any situation. In order to make the above definition of the subtyping relation syntax-directed, the transitivity rule \( SA-trans-TVar \) is restricted to the case where the type on the left-hand side of the \(<\) operator is a variable. However, the unrestricted transitivity rule can be derived from this definition. In order for the proof to go through, we have to simultaneously prove another property called narrowing. The two properties are proved by nested induction. The outer induction is on the size of the type \( Q \), whereas the two inner inductions for proving transitivity and narrowing are on the derivation of the subtyping judgements. The transitivity property is needed in the proof of narrowing, which is by induction on the derivation of \( \Delta @ TVarB \ 0 P :: \Gamma \vdash \ 0 Q \). In the case corresponding to the rule \( SA-trans-TVar \), we must prove \( \Delta @ TVarB \ 0 P :: \Gamma \vdash \ 0 Q \) i \(<\): T. The only interesting case is the one where \( i = \| \Delta \| \). By induction hypothesis, we know that \( \Delta @ TVarB \ 0 P :: \Gamma \vdash \uparrow_\tau (i + 1) \ 0 Q :: T \) and \( \Delta @ TVarB \ Q :: \Gamma \langle i \rangle = \{ TVarB \ 0 Q \} \). By assumption, we have \( \Gamma \vdash P <: Q \) and hence \( \Delta @ TVarB \ 0 P :: \Gamma \vdash \uparrow_\tau (i + 1) \ 0 P :: \Gamma \vdash \uparrow_\tau (i + 1) \ 0 Q \) by weakening. Since \( \uparrow_\tau (i + 1) \ 0 Q \) has the same size as \( Q \), we can use the transitivity property, which yields \( \Delta @ TVarB \ 0 P :: \Gamma \vdash \uparrow_\tau (i + 1) \ 0 P :: T \). The claim then follows easily by an application of \( SA-trans-TVar \).

**Lemma subtype-trans: — A.3**

\( \Gamma \vdash S <: Q \Rightarrow \Gamma \vdash Q <: T \Rightarrow \Gamma \vdash S <: T \)

\( \Delta @ TVarB \ Q :: \Gamma \vdash M <: N \Rightarrow \Gamma \vdash P <: Q \Rightarrow \Delta @ TVarB \ P :: \Gamma \vdash M <: N \)

In the proof of the preservation theorem presented in §2.6, we will also need a substitution theorem, which is proved by induction on the subtyping derivation:

11
lemma substT-subtype: \[\Delta @ TVarB Q :: \Gamma \vdash S <: T\]
shows \[\Gamma \vdash P <: Q \implies \Delta[0 \mapsto \tau] P @ \Gamma \vdash S[\parallel \Delta \parallel \mapsto \tau] P : T[\parallel \Delta \parallel \mapsto \tau] P]\n(proof)

lemma subst-subtype:
assumes \[H: \Delta @ VarB V :: \Gamma \vdash T <: U\]
shows \[\downarrow_{\tau} t 0 \Delta @ \Gamma \vdash \downarrow_{\tau} T \parallel \Delta \parallel T <: \downarrow_{\tau} \parallel \Delta \parallel U\]
(proof)

2.5 Typing

We are now ready to give a definition of the typing judgement \(\Gamma \vdash t : T\).

inductive typing :: env \Rightarrow trm \Rightarrow type \Rightarrow bool \quad (\vdash \vdash : [50, 50, 50])
where
| T-Var: \(\Gamma \vdash_w t \implies \Gamma(i) = [VarB U] \implies T = \uparrow \tau (Suc i) 0 U \implies \Gamma \vdash Var i : T\)
| T-Abs: VarB T1 :: \(\Gamma \vdash t_2 : T_2 \implies \Gamma \vdash (\lambda : T_1. t_2) : T_1 \rightarrow T_0 T_2\)
| T-App: \(\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \implies \Gamma \vdash t_2 : T_{11} \implies \Gamma \vdash t_1 \cdot t_2 : T_{12}\)
| T-TAbs: TVarB T1 :: \(\Gamma \vdash t_2 : T_2 \implies \Gamma \vdash (\forall : T_1, t_2) : (\forall : T_1, T_2)\)
| T-TApp: \(\Gamma \vdash t_1 : (\forall : T_{11}, T_{12}) \implies \Gamma \vdash T_2 <: T_{11} \implies \Gamma \vdash t_1 \cdot t_2 : T_{12}[0 \mapsto \tau] T_{12}[\tau]\)
| T-Sub: \(\Gamma \vdash t : S \implies \Gamma \vdash S <: T \implies \Gamma \vdash t : T\)

Note that in the rule T-Var, the indices of the type \(U\) looked up in the context \(\Gamma\) need to be incremented in order for the type to be well-formed with respect to \(\Gamma\). In the rule T-Abs, the type \(T_2\) of the abstraction body \(t_2\) may not contain the variable with index \(0\), since it is a term variable. To compensate for the disappearance of the context element VarB \(T_1\) in the conclusion of thy typing rule, the indices of all free type variables in \(T_2\) must be decremented by \(1\).

theorem wf-typeE1:
assumes \(H: \Gamma \vdash t : T\)
shows \(\Gamma \vdash_w t\) (proof)

theorem wf-typeE2:
assumes \(H: \Gamma \vdash t : T\)
shows \(\Gamma \vdash_w t\) (proof)

Like for the subtyping judgement, we can again prove that all types and contexts involved in a typing judgement are well-formed:

lemma wf-type-conj: \(\Gamma \vdash t : T \implies \Gamma \vdash_w t \land \Gamma \vdash_w t\)
(proof)

The narrowing theorem for the typing judgement states that replacing the type of a variable in the context by a subtype preserves typability:

lemma narrow-type: \(\Delta @ TVarB Q :: \Gamma \vdash S <: T\)
(proof)

12
assumes $H$: $\Delta \Atoms TVarB Q \because \Gamma \vdash t : T$
shows $\Gamma \vdash P <: Q \Longrightarrow \Delta \Atoms TVarB P \because \Gamma \vdash t : T$
(proof)

lemma subtype-refl':
assumes $t: \Gamma \vdash t : T$
shows $\Gamma \vdash T <: T$
(proof)

lemma $Abs$-type: — A.13(1)
assumes $H: \Gamma \vdash (\lambda S.s) : T$
shows $\Gamma \vdash T <: U \rightarrow U' \Longrightarrow$
$$\begin{align*}
(\forall S'. \Gamma \vdash U <: S \Longrightarrow VarB S :: \Gamma \vdash s : S' \Longrightarrow \\
\Gamma \vdash \downarrow 0 S' <: U' \Longrightarrow P) \Longrightarrow P
\end{align*}$$
(proof)

lemma $Abs$-type':
assumes $H: \Gamma \vdash (\lambda S.s) : U \rightarrow U'$
and $R: \forall S'. \Gamma \vdash U <: S \Longrightarrow VarB S :: \Gamma \vdash s : S' \Longrightarrow \\
\Gamma \vdash \downarrow 0 S' <: U' \Longrightarrow P$
shows $P$ (proof)

lemma $TAbs$-type: — A.13(2)
assumes $H: \Gamma \vdash (\lambda < S.s) : T$
shows $\Gamma \vdash T <: (\forall < U. U') \Longrightarrow$
$$\begin{align*}
(\forall S'. \Gamma \vdash U <: S \Longrightarrow TVarB U :: \Gamma \vdash s : S' \Longrightarrow \\
TVarB U :: \Gamma \vdash S' <: U' \Longrightarrow P) \Longrightarrow P
\end{align*}$$
(proof)

lemma $TAbs$-type':
assumes $H: \Gamma \vdash (\lambda < S.s) : (\forall < U. U')$
and $R: \forall S'. \Gamma \vdash U <: S \Longrightarrow TVarB U :: \Gamma \vdash s : S' \Longrightarrow \\
TVarB U :: \Gamma \vdash S' <: U' \Longrightarrow P$
shows $P$ (proof)

lemma $T$-eq: $\Gamma \vdash t : T \Longrightarrow T = T' \Longrightarrow \Gamma \vdash t : T'$ (proof)

The weakening theorem states that inserting a binding $B$ does not affect typing:

lemma type-weaken:
assumes $H: \Delta \Atoms t : T$
shows $\Gamma \vdash wfB B \Longrightarrow$
$$\begin{align*}
\uparrow_\tau 0 \Delta \Atoms B :: \Gamma \vdash \uparrow I \parallel \Delta \parallel t : \uparrow_\tau I \parallel \Delta \parallel T
\end{align*}$$
(proof)

We can strengthen this result, so as to mean that concatenating a new context $\Delta$ to the context $\Gamma$ preserves typing:

lemma type-weaken': — A.5(6)
$$\Gamma \vdash t : T \Longrightarrow \Delta \Atoms \Gamma \vdash t : T \Longrightarrow \Delta \Atoms \Gamma \vdash \Delta \Atoms t : T$$
(proof)
This property is proved by structural induction on the context $\Delta$, using the previous result in the induction step. In the proof of the preservation theorem, we will need two substitution theorems for term and type variables, both of which are proved by induction on the typing derivation. Since term and type variables are stored in the same context, we again have to decrement the free type variables in $\Delta$ and $T$ by 1 in the substitution rule for term variables in order to compensate for the disappearance of the variable.

**Theorem subst-type:** — \(\Lambda 8\)

**Assumes**

\(H: \Delta @ VarB U :: \Gamma \vdash t : T\)

**Shows** \(\Gamma \vdash u : U \implies I \Gamma \vdash t[\|\Delta\| \mapsto \|\Delta\|] : T\) (proof)

**Theorem substT-type:** — \(\Lambda 11\)

**Assumes**

\(H: \Delta @ TVarB Q :: \Gamma \vdash t : T\)

**Shows** \(\Gamma \vdash P < Q \implies \Delta[0 \mapsto \|\Delta\| \mapsto \|\Delta\|] \vdash t[\|\Delta\| \mapsto \|\Delta\|] : T[\|\Delta\| \mapsto \|\Delta\|] \) (proof)

### 2.6 Evaluation

For the formalization of the evaluation strategy, it is useful to first define a set of canonical values that are not evaluated any further. The canonical values of call-by-value $F_{<}$ are exactly the abstractions over term and type variables:

**Inductive-set**

\[\text{value} :: \text{trm set}\]

**Where**

\(\text{Abs}: (\lambda:T. t) \in \text{value}\)
\(\text{TAbs}: (\lambda<:T. t) \in \text{value}\)

The notion of a value is now used in the definition of the evaluation relation $t \mapsto t'$. There are several ways for defining this evaluation relation: Aydemir et al. [1] advocate the use of evaluation contexts that allow to separate the description of the "immediate" reduction rules, i.e. $\beta$-reduction, from the description of the context in which these reductions may occur in. The rationale behind this approach is to keep the formalization more modular. We will take a closer look at this style of presentation in section §4. For the rest of this section, we will use a different approach: both the "immediate" reductions and the reduction context are described within the same inductive definition, where the context is described by additional congruence rules.

**Inductive**

\[\text{eval} :: \text{trm} \Rightarrow \text{trm} \Rightarrow \text{bool} \ (\text{infixl} \mapsto 50)\]

**Where**

\(E-\text{Abs}: v_2 \in \text{value} \implies (\lambda:T_{11}. t_{12}) \cdot v_2 \mapsto t_{12}[0 \mapsto v_2]\)
\(E-\text{TAbs}: (\lambda<:T_{11}. t_{12}) \cdot \tau \mapsto t_{12}[0 \mapsto \tau \ T_{2}]\)
\(E-\text{App1}: t \mapsto t' \implies t \cdot u \mapsto t' \cdot u\)
\(E-\text{App2}: v \in \text{value} \implies t \mapsto t' \implies v \cdot t \mapsto v \cdot t'\)
Here, the rules $E\text{-Abs}$ and $E\text{-TAbs}$ describe the “immediate” reductions, whereas $E\text{-App1}$, $E\text{-App2}$, and $E\text{-TApp}$ are additional congruence rules describing reductions in a context. The most important theorems of this section are the preservation theorem, stating that the reduction of a well-typed term does not change its type, and the progress theorem, stating that reduction of a well-typed term does not “get stuck” – in other words, every well-typed, closed term $t$ is either a value, or there is a term $t'$ to which $t$ can be reduced. The preservation theorem is proved by induction on the derivation of $\Gamma \vdash t : T$, followed by a case distinction on the last rule used in the derivation of $t \longrightarrow t'$.

**Theorem** Preservation: — A.20

assumes $H: \Gamma \vdash t : T$

shows $t \longrightarrow t' \Longrightarrow \Gamma \vdash t' : T$  \(\langle \text{proof} \rangle\)

The progress theorem is also proved by induction on the derivation of $\parallel \vdash t : T$. In the induction steps, we need the following two lemmas about canonical forms stating that closed values of types $T_1 \rightarrow T_2$ and $\forall \ll T_1$. $T_2$ must be abstractions over term and type variables, respectively.

**Lemma** Fun-canonical: — A.14(1)

assumes $ty: \parallel \vdash v : T_1 \rightarrow T_2$

shows $v \in value \Longrightarrow \exists t. S. v = (\lambda S. t)$  \(\langle \text{proof} \rangle\)

**Lemma** TyAll-canonical: — A.14(3)

assumes $ty: \parallel \vdash v : (\forall \ll T_1. T_2)$

shows $v \in value \Longrightarrow \exists t. S. v = (\lambda S. t)$  \(\langle \text{proof} \rangle\)

**Theorem** Progress:

assumes $ty: \parallel \vdash t : T$

shows $t \in value \lor (\exists t'. t \longrightarrow t')$  \(\langle \text{proof} \rangle\)

### 3 Extending the calculus with records

We now describe how the calculus introduced in the previous section can be extended with records. An important point to note is that many of the definitions and proofs developed for the simple calculus can be reused.

#### 3.1 Types and Terms

In order to represent records, we also need a type of field names. For this purpose, we simply use the type of strings. We extend the datatype of types of System $\mathbb{F}_\ll$ by a new constructor $RcdT$ representing record types.

**Type-synonym** name = string
datatype type =
  TVar nat
| Top
| Fun type type  (infixr \to 200)
| TyAll type type ((\forall <:\cdot \cdot \cdot \cdot[0, 10] 10)
| RcdT (name \times type) list

type-synonym fldT = name \times type

type-synonym rcdT = (name \times type) list

datatype binding = VarB type | TVarB type

type-synonym env = binding list

primrec is-TVarB :: binding => bool
where
  is-TVarB (VarB T) = False
| is-TVarB (TVarB T) = True

primrec type-ofB :: binding => type
where
  type-ofB (VarB T) = T
| type-ofB (TVarB T) = T

primrec mapB :: (type => type) => binding => binding
where
  mapB f (VarB T) = VarB (f T)
| mapB f (TVarB T) = TVarB (f T)

A record type is essentially an association list, mapping names of record fields to their types. The types of bindings and environments remain unchanged. The datatype \textit{trm} of terms is extended with three new constructors \textit{Rcd}, \textit{Proj}, and \textit{LET}, denoting construction of a new record, selection of a specific field of a record (projection), and matching of a record against a pattern, respectively. A pattern, represented by datatype \textit{pat}, can be either a variable matching any value of a given type, or a nested record pattern. Due to the encoding of variables using de Bruijn indices, a variable pattern only consists of a type.

datatype pat = PVar type | PRcd (name \times pat) list

datatype trm =
  Var nat
| Abs type trm  ((\lambda <:\cdot \cdot \cdot \cdot[0, 10] 10)
| TAbs type trm ((\lambda <:\cdot \cdot \cdot \cdot[0, 10] 10)
| App trm trm  (infixl . 200)
| TApp trm type (infixl . \cdot 200)
| Rcd (name \times trm) list
In order to motivate the typing and evaluation rules for the \( \text{LET} \), it is important to note that an expression of the form
\[
\text{LET} \ PRcd \ [(l_1, \ PVar \ T_1), \ldots, (l_n, \ PVar \ T_n)] = Rcd \ [(l_1, \ v_1), \ldots, (l_n, \ v_n)] \ IN \ t
\]
can be treated like a nested abstraction \((\lambda: T_1. \ldots \lambda: T_n. \ t) \cdot v_1 \cdot \ldots \cdot v_n\)

### 3.2 Lifting and Substitution

**primrec** \( \text{psize} :: \ p \Rightarrow \ \text{nat} (\|\cdot\|_p) \)

and \( \text{rsize} :: \ rpat \Rightarrow \ \text{nat} (\|\cdot\|_r) \)

and \( \text{fsize} :: \ fpat \Rightarrow \ \text{nat} (\|\cdot\|_f) \)

**where**

\[
\|PVar \ T\|_p = 1
\]

\[
\|PRcd \ fs\|_p = \|fs\|_r
\]

\[
\|P\|_r = 0
\]

\[
\|f :: \ fs\|_r = \|f\|_f + \|fs\|_r
\]

\[
\|(l, \ p)\|_f = \|p\|_p
\]

**primrec** \( \text{liftT} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{type} \Rightarrow \ \text{type} (\uparrow_\tau) \)

and \( \text{liftrT} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{rcdT} \Rightarrow \ \text{rcdT} (\uparrow_\tau) \)

and \( \text{liftfT} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{fldT} \Rightarrow \ \text{fldT} (\uparrow_\tau) \)

**where**

\[
\uparrow_\tau \ n \ k \ (TVar \ i) = (\text{if } i < k \ \text{then } TVar \ i \ \text{else } TVar \ (i + n))
\]

\[
\uparrow_\tau \ n \ k \ (T \rightarrow U) = \uparrow_\tau \ n \ k \ T \rightarrow \uparrow_\tau \ n \ k \ U
\]

\[
\uparrow_\tau \ n \ k \ (\forall <T. \ U) = (\forall <\uparrow_\tau \ n \ k \ T. \ \uparrow_\tau \ n \ (k + 1) \ U)
\]

\[
\uparrow_\tau \ n \ k \ (RcdT \ fs) = RcdT (\uparrow_\tau \ n \ k \ fs)
\]

\[
\uparrow_\tau \ n \ k \ [] = []
\]

\[
\uparrow_\tau \ n \ k \ (f :: \ fs) = \uparrow_\tau \ n \ k \ f :: \uparrow_\tau \ n \ k \ fs
\]

\[
\uparrow_\tau \ n \ k \ (l, \ T) = (l, \uparrow_\tau \ n \ k \ T)
\]

**primrec** \( \text{liftp} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{pat} \Rightarrow \ \text{pat} (\uparrow_p) \)

and \( \text{liftrp} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{rpat} \Rightarrow \ \text{rpat} (\uparrow_p) \)

and \( \text{liftfp} :: \ \text{nat} \Rightarrow \ \text{nat} \Rightarrow \ \text{fpat} \Rightarrow \ \text{fpat} (\uparrow_p) \)

**where**

\[
\uparrow_p \ n \ k \ (PVar \ T) = PVar (\uparrow_p \ n \ k \ T)
\]

\[
\uparrow_p \ n \ k \ (PRcd \ fs) = PRcd (\uparrow_p \ n \ k \ fs)
\]

\[
\uparrow_p \ n \ k \ [] = []
\]

\[
\uparrow_p \ n \ k \ (f :: \ fs) = \uparrow_p \ n \ k \ f :: \uparrow_p \ n \ k \ fs
\]

\[
\uparrow_p \ n \ k \ (l, \ p) = (l, \uparrow_p \ n \ k \ p)
\]
patterns, which we denote by ↑

basic calculus, we also have to define lifting and substitution functions for

In addition to the lifting and substitution functions already needed for the
decp

|↓

where

|↑ n k (Var i) = (if i < k then Var i else Var (i + n))
|↑ n k (λ::T. t) = (λ↑↑↑↑↑ n k T. ↑↑↑↑↑ n (k + 1) t)
|↑ n k (λ<s::T. t) = (λ<s::T↑↑ n k T. ↑↑ n (k + 1) t)
|↑ n k (s · t) = ↑↑↑↑↑ n k s · ↑↑ n k t
|↑ n k (t • T) = ↑↑↑↑↑ n k t • T↑↑ n k T
|↑ n k (Rcd fs) = Rcd (↑↑↑↑↑ n k fs)
|↑ n k (t ..a) = (↑↑↑↑↑ n k t) .a
|↑↑↑↑↑ n k (λ f : fs) = ↑↑↑↑↑ n k f : ↑↑↑↑↑ n k fs
|↑↑↑↑↑ n k (l, t) = (l, ↑↑↑↑↑ n k t)

primrec lift :: nat ⇒ nat ⇒ trm ⇒ trm (↑)
and liftr :: nat ⇒ nat ⇒ rcd ⇒ rcd (↑r)
and liffr :: nat ⇒ nat ⇒ fld ⇒ fld (↑f)
where

| n k (Var i) = (if i < k then Var i else Var (i + n))
| n k (λ::T. t) = (λ↑↑↑↑↑ n k T. ↑↑↑↑↑ n (k + 1) t)
| n k (λ<s::T. t) = (λ<s::T↑↑ n k T. ↑↑ n (k + 1) t)
| n k (s · t) = ↑↑↑↑↑ n k s · ↑↑ n k t
| n k (t • T) = ↑↑↑↑↑ n k t • T↑↑ n k T
| n k (Rcd fs) = Rcd (↑↑↑↑↑ n k fs)
| n k (t ..a) = (↑↑↑↑↑ n k t) .a
| (λ f : fs) = ↑↑↑↑↑ n k f : ↑↑↑↑↑ n k fs
| (l, t) = (l, ↑↑↑↑↑ n k t)

primrec substTT :: type ⇒ nat ⇒ type ⇒ type (-[·⇒τ · ]τ [300, 0, 0] 300)
and substTT :: rcdT ⇒ nat ⇒ type ⇒ rcdT (-[·⇒τ · ]τ [300, 0, 0] 300)
and substfTT :: fldT ⇒ nat ⇒ type ⇒ fldT (-[·⇒τ · ]τ [300, 0, 0] 300)
where

(TVar i)[k ⇒ τ] S τ =
(if k < i then TVar (i - 1) else if i = k then ↑↑↑↑↑ k 0 S else TVar i)
| Top [k ⇒ τ] S τ = Top
| (T ⇒ U)[k ⇒ τ] S τ = (T[k ⇒ τ] S τ ⇒ U[k ⇒ τ] S τ)
| (∀·: T. U)[k ⇒ τ] S τ = (∀·: T[k ⇒ τ] S τ. U[k+1 ⇒ τ] S τ)
| (Rcd fs)[k ⇒ τ] S τ = Rcd (fs[k ⇒ τ] S τ)
| [] [k ⇒ τ] S τ = []
| (λ f : fs)[k ⇒ τ] S τ = f[k ⇒ τ] S f τ :: fs[k ⇒ τ] S f τ
| (l, T)[k ⇒ τ] S f τ = (l, T[k ⇒ τ] S f τ)

primrec substpT :: pat ⇒ nat ⇒ type ⇒ pat (-[·⇒τ · ]τ [300, 0, 0] 300)
and substpTT :: rpat ⇒ nat ⇒ type ⇒ rpat (-[·⇒τ · ]τ [300, 0, 0] 300)
and substfT :: fpat ⇒ nat ⇒ type ⇒ fpat (-[·⇒τ · ]τ [300, 0, 0] 300)
where

(PVar T)[k ⇒ τ] S p = PVar (T[k ⇒ τ] S p)
| (PRcd fs)[k ⇒ τ] S p = PRcd (fs[k ⇒ τ] S f p)
| [] [k ⇒ τ] S f p = []
| (λ f : fs)[k ⇒ τ] S f p = f[k ⇒ τ] S f f p :: fs[k ⇒ τ] S f f p
| (l, p)[k ⇒ τ] S f f p = (l, p[k ⇒ τ] S p)

primrec depc :: nat ⇒ nat ⇒ pat ⇒ pat (↓p)
where

↓p 0 k p = p
| ↓p (Suc n) k p = ↓p n k (p[k ⇒ τ] Top p)

In addition to the lifting and substitution functions already needed for the

basic calculus, we also have to define lifting and substitution functions for

patterns, which we denote by ↑ p n k p and T[k ⇒ τ] S τ, respectively. The
extension of the existing lifting and substitution functions to records is fairly standard.

**primrec subst :: trm \Rightarrow nat \Rightarrow trm (\cdot \Rightarrow \cdot) [300, 0, 0] 300**

**and subst :: rec \Rightarrow nat \Rightarrow trm \Rightarrow rec (\cdot \Rightarrow \cdot)_r [300, 0, 0] 300**

**and substf :: fld \Rightarrow nat \Rightarrow trm \Rightarrow fld (\cdot \Rightarrow \cdot_f) [300, 0, 0] 300**

where

\[(\text{Var } i)[k \mapsto s] =\]

\[(\text{if } k < i \text{ then } \text{Var } (i - 1) \text{ else if } i = k \text{ then } \uparrow k 0 s \text{ else } \text{Var } i)\]

\[(t \cdot u)[k \mapsto s] = t[k \mapsto s] \cdot u[k \mapsto s]\]

\[(t \cdot T)[k \mapsto s] = t[k \mapsto s] \cdot T[k \mapsto \text{Top}]_\tau\]

\[(\lambda T \cdot t)[k \mapsto s] = (\lambda T[k \mapsto \text{Top}]_\tau \cdot t[k \mapsto s])\]

\[(\lambda T. t)[k \mapsto s] = (\lambda T[k \mapsto \text{Top}]_\tau \cdot t[k \mapsto s])\]

\[(\text{Red } fs)[k \mapsto s] = \text{Red } (fs[k \mapsto s])_\tau\]

\[(t \cdot a)[k \mapsto s] = (t[k \mapsto s])_\cdot a\]

\[(\text{LET } p = t \text{ IN } u)[k \mapsto s] = (\text{LET } \downarrow p \text{ IN } u)[k \mapsto s] (\text{LET } \downarrow p \text{ IN } u[k + \|p\|_p \mapsto s])\]

\[\text{Note that the substitution function on terms is defined simultaneously with a substitution function } fs[k \mapsto s]_\tau \text{ on records (i.e. lists of fields), and a substitution function } f[k \mapsto s]_f \text{ on fields. To avoid conflicts with locally bound variables, we have to add an offset } \|p\|_p \text{ to } k \text{ when performing substitution in the body of the LET binder, where } \|p\|_p \text{ is the number of variables in the pattern } p.\]

**primrec substT :: trm \Rightarrow type \Rightarrow trm (\cdot \Rightarrow \cdot) [300, 0, 0] 300**

**and substT :: rec \Rightarrow nat \Rightarrow type \Rightarrow rec (\cdot \Rightarrow \cdot)_r [300, 0, 0] 300**

**and substfT :: fld \Rightarrow nat \Rightarrow type \Rightarrow fld (\cdot \Rightarrow \cdot_f) [300, 0, 0] 300**

where

\[(\text{Var } i)[k \mapsto s] = (\text{if } k < i \text{ then } \text{Var } (i - 1) \text{ else } \text{Var } i)\]

\[(t \cdot u)[k \mapsto s] = t[k \mapsto s] \cdot u[k \mapsto s]\]

\[(t \cdot T)[k \mapsto s] = t[k \mapsto s] \cdot T[k \mapsto \text{Top}]_\tau\]

\[(\lambda T. t)[k \mapsto s] = (\lambda T[k \mapsto \text{Top}]_\tau \cdot t[k \mapsto s])\]

\[(\lambda T. t)[k \mapsto s] = (\lambda T[k \mapsto \text{Top}]_\tau \cdot t[k \mapsto s])\]

\[(\text{Red } fs)[k \mapsto s] = \text{Red } (fs[k \mapsto s])_\tau\]

\[(t \cdot a)[k \mapsto s] = (t[k \mapsto s])_\cdot a\]

\[(\text{LET } p = t \text{ IN } u)[k \mapsto s] =\]

\[(\text{LET } p = t \text{ IN } u)[k \mapsto s] = (\text{LET } \downarrow p \text{ IN } u)[k \mapsto s] (\text{LET } \downarrow p \text{ IN } u[k + \|p\|_p \mapsto s])\]

\[\text{primrec liftE :: nat } \Rightarrow \text{ nat } \Rightarrow \text{ env } \Rightarrow \text{ env } (\uparrow_\varepsilon)\]

**where**

\[\uparrow_\varepsilon n k 0 \cdot = 0\]

\[\uparrow_\varepsilon n k (B :: \Gamma) = \text{mapB } (\uparrow_\varepsilon n (k + \|\Gamma\|)) B :: \uparrow_\varepsilon n k \Gamma\]

**primrec substE :: env \Rightarrow nat \Rightarrow type \Rightarrow env (\cdot \Rightarrow \cdot)_e [300, 0, 0] 300**
For the formalization of the reduction rules for LET, we need a function 
\( t[k \mapsto_s us] \) for simultaneously substituting terms \( us \) for variables with 
consecutive indices:

```plaintext
primrec subst :: trm \Rightarrow nat \Rightarrow trm \Rightarrow trm \ (\cdot \mapsto_s \cdot) \ [300, 0, 0] 300
where
|\\langle t, [], t[k \mapsto_s us] \rangle = t
|\\langle t, [u \mapsto u], t[k \mapsto_s us] \rangle = t[k + ||us|| \mapsto u] [k \mapsto_s us]
```

```plaintext
primrec dec :: nat \Rightarrow nat \Rightarrow type \Rightarrow type (↓)
where
↓n 0 \ k T = T
|↓n (Suc n) k T = ↓n k (T[k \mapsto_T Top]r)
```

```plaintext
primrec decE :: nat \Rightarrow nat \Rightarrow env \Rightarrow env (\downarrow_e)
where
↓e 0 \ k \ Γ = \ Γ
|↓e (Suc n) k \ Γ = ↓e n k (Γ[k \mapsto_T Top]e)
```

```plaintext
primrec decr :: nat \Rightarrow nat \Rightarrow rcd \Rightarrow rcd (↓_r)
where
↓r 0 \ fTs = fTs
|↓r (Suc n) \ fTs = ↓r n k (fTs[k \mapsto_T Top]r)
```

The lemmas about substitution and lifting are very similar to those needed 
for the simple calculus without records, with the difference that most of 
them have to be proved simultaneously with a suitable property for records.

```plaintext
lemma liftE-length [simp]: ||↑n k \ Γ|| = ||Γ||
(proof)
```

```plaintext
lemma substE-length [simp]: ||Γ[k \mapsto_T U]|e|| = ||Γ||
(proof)
```

```plaintext
lemma liftE-nth [simp]:
(↑n k \ Γ)(\langle i \rangle) = map-option (mapB (↑n (k + ||Γ|| − i − 1))) (Γ\langle i \rangle)
(proof)
```

```plaintext
lemma substE-nth [simp]:
(Γ[0 \mapsto_s T]|e)(\langle i \rangle) = map-option (mapB (λU. U[||Γ|| − i − 1 \mapsto_T T]r) (Γ\langle i \rangle)
(proof)
```

```plaintext
lemma liftT-liftT [simp]:
i ≤ j \Rightarrow j ≤ i + m \Rightarrow ↑n j \ (↑n m i T) = ↑n (m + n) i T
i ≤ j \Rightarrow j ≤ i + m \Rightarrow ↑r j \ (↑r m i rT) = ↑r (m + n) i rT
i ≤ j \Rightarrow j ≤ i + m \Rightarrow ↑f j \ (↑f m i fT) = ↑f (m + n) i fT
(proof)
```

20
lemma liftT-liftT' [simp]:
\[ i + m \leq j \implies \uparrow_{\tau} n j (\uparrow_{\tau} m i T) = \uparrow_{\tau} m i (\uparrow_{\tau} n (j - m) T) \]
\[ i + m \leq j \implies \uparrow_{\tau T} n j (\uparrow_{\tau T} m i rT) = \uparrow_{\tau T} m i (\uparrow_{\tau T} n (j - m) rT) \]
\[ i + m \leq j \implies \uparrow_{f T} n j (\uparrow_{f T} m i fT) = \uparrow_{f T} m i (\uparrow_{f T} n (j - m) fT) \]
(proof)

lemma lift-size [simp]:
\[ \text{size} (\uparrow_{\tau} n k T) = \text{size} T \]
\[ \text{size-list} (\text{size-prod} (\lambda x. 0) \text{size}) (\uparrow_{\tau T} n k rT) = \text{size-list} (\text{size-prod} (\lambda x. 0) \text{size}) rT \]
\[ \text{size-prod} (\lambda x. 0) \text{size} (\uparrow_{f T} n k fT) = \text{size-prod} (\lambda x. 0) \text{size} fT \]
(proof)

lemma liftT0 [simp]:
\[ \uparrow_{\tau} 0 i T = T \]
\[ \uparrow_{\tau T} 0 i rT = rT \]
\[ \uparrow_{f T} 0 i fT = fT \]
(proof)

lemma liftp0 [simp]:
\[ \uparrow_{\tau} 0 i p = p \]
\[ \uparrow_{\tau p} 0 i f s = f s \]
\[ \uparrow_{f p} 0 i f = f \]
(proof)

lemma lift0 [simp]:
\[ \uparrow 0 i t = t \]
\[ \uparrow_{\tau} 0 i f s = f s \]
\[ \uparrow_{f} 0 i f = f \]
(proof)

theorem substT-liftT [simp]:
\[ k \leq k' \implies k' < k + n \implies (\uparrow_{\tau} n k T)[k' \mapsto_{\tau} U]_{\tau} = \uparrow_{\tau} (n - 1) k T \]
\[ k \leq k' \implies k' < k + n \implies (\uparrow_{\tau T} n k rT)[k' \mapsto_{\tau T} U]_{\tau T} = \uparrow_{\tau T} (n - 1) k rT \]
\[ k \leq k' \implies k' < k + n \implies (\uparrow_{f T} n k fT)[k' \mapsto_{f T} U]_{f T} = \uparrow_{f T} (n - 1) k fT \]
(proof)

theorem liftT-substT [simp]:
\[ k \leq k' \implies \uparrow_{\tau} n k (T[k' \mapsto_{\tau} U]_{\tau}) = \uparrow_{\tau} n k T[k' + n \mapsto_{\tau} U]_{\tau} \]
\[ k \leq k' \implies \uparrow_{\tau T} n k (rT[k' \mapsto_{\tau T} U]_{\tau T}) = \uparrow_{\tau T} n k rT[k' + n \mapsto_{\tau T} U]_{\tau T} \]
\[ k \leq k' \implies \uparrow_{f T} n k (fT[k' \mapsto_{f T} U]_{f T}) = \uparrow_{f T} n k fT[k' + n \mapsto_{\tau T} U]_{f T} \]
(proof)

theorem liftT-substT' [simp]:
\[ k' < k \implies \uparrow_{\tau} n k (T[k' \mapsto_{\tau} U]_{\tau}) = \uparrow_{\tau} n (k + 1) T[k' \mapsto_{\tau} U]_{\tau} \]
\[ k' < k \implies \uparrow_{\tau T} n k (rT[k' \mapsto_{\tau T} U]_{\tau T}) = \uparrow_{\tau T} n (k + 1) rT[k' \mapsto_{\tau T} U]_{\tau T} \]

\[ k' < k \implies \uparrow_{f T} n k (fT[k' \mapsto_{f T} U]_{f T}) = \uparrow_{f T} n (k + 1) fT[k' \mapsto_{\tau T} U]_{f T} \]
\[ k' < k \implies \]
\[ \uparrow n \ (T[k' \mapsto \tau] T_\tau) = \uparrow n \ (k + 1) \] \[ (k + 1) \] \[ T[k' \mapsto \tau] T_\tau \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-substT-Top [simp]:
\[ k \leq k' \implies \uparrow n \ k' \ (T[k \mapsto \tau] Top_\tau) = \uparrow n \ (Succ k') \ T[k \mapsto \tau] Top_\tau \]
\[ k \leq k' \implies \uparrow \tau \ n \ k' \ (rT[k' \mapsto \tau] Top_{\tau\tau}) = \uparrow \tau \ n \ (Succ k') \ rT[k \mapsto \tau] Top_{\tau\tau} \]
\[ k \leq k' \implies \uparrow n \ k' \ (T[k \mapsto \tau] Top_{\tau\tau}) = \uparrow n \ (Succ k') \ T[k \mapsto \tau] Top_{\tau\tau} \]
\[ \langle \text{proof} \rangle \]

**Theorem** liftE-substE [simp]:
\[ k' \leq k \implies \uparrow n \ k' \ (\Gamma[k' \mapsto \tau] U_\tau) = \uparrow n \ k \ (\Gamma + n \mapsto \tau) U_\tau \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-decT [simp]:
\[ k' \leq k \implies \uparrow n \ k' \ (\downarrow n k m \ k T) = \downarrow n k (\uparrow n (m + k')) T \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-substT-strange:
\[ \uparrow n \ k \ T[n + k \mapsto \tau] U_\tau = \uparrow n \ (Succ k) \ T[k \mapsto \tau] \uparrow n \ 0 \ U_\tau \]
\[ \uparrow n \ k \ rT[n + k \mapsto \tau] U_{\tau\tau} = \uparrow n \ (Succ k) \ rT[k \mapsto \tau] \uparrow n \ 0 \ U_{\tau\tau} \]
\[ \uparrow n \ k \ fT[n + k \mapsto \tau] U_f\tau = \uparrow n \ (Succ k) \ fT[k \mapsto \tau] \uparrow n \ 0 \ U_f\tau \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-liftp [simp]:
\[ k' \leq k + n \implies \uparrow n \ k' \ (\uparrow n k p) = \uparrow n' \ (n + k') \ p \]
\[ k' \leq k + n \implies \uparrow n \ k' \ (\uparrow n k p) = \uparrow n' \ (n + k') \ p \]
\[ k' \leq k + n \implies \uparrow n \ k' \ (\uparrow n k f) = \uparrow n' \ (n + k') \ f \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-liftp-psize [simp]:
\[ \| \uparrow n k p \|_p = \| p \|_p \]
\[ \| \uparrow n k f \|_f = \| f \|_f \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-lift [simp]:
\[ k \leq k' \implies k' \leq k + n \implies \uparrow n' k' \ (\uparrow n k t) = \uparrow n' (n + k') \ t \]
\[ k \leq k' \implies k' \leq k + n \implies \uparrow n' k' \ (\uparrow n k f) = \uparrow n' (n + k') \ f \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-liftE [simp]:
\[ k \leq k' \implies k' \leq k + n \implies \uparrow n' k' \ (\uparrow n k \Gamma) = \uparrow n' (n + k') \ k \Gamma \]
\[ \langle \text{proof} \rangle \]

**Lemma** liftF-liftE' [simp]:
\[ i + m \leq j \implies \uparrow n j \ (\uparrow n m i \ Gamma) = \uparrow n (j - m) \]
\[ \langle \text{proof} \rangle \]
lemma substT-substT:
\[ i \leq j \implies T[Suc j \mapsto \tau V]\tau[i \mapsto \tau U]\tau[j - i \mapsto \tau V]\tau = T[i \mapsto \tau U]\tau[j \mapsto \tau V]\tau \]
\[ i \leq j \implies rT[Suc j \mapsto \tau V]\tau[i \mapsto \tau U]\tau[j - i \mapsto \tau V]\tau = rT[i \mapsto \tau U]\tau[j \mapsto \tau V]\tau \]
\[ i \leq j \implies fT[Suc j \mapsto \tau V]\tau[f\tau[i \mapsto \tau U]\tau[j - i \mapsto \tau V]\tau = fT[i \mapsto \tau U]\tau[f\tau[j \mapsto \tau V]\tau \]
(\text{proof})

lemma substT-decT [simp]:
\[ k \leq j \implies (\downarrow i k T)[j \mapsto \tau U]\tau = \downarrow i k (T[i + j \mapsto \tau U]\tau) \]
(\text{proof})

lemma substT-decT' [simp]:
\[ i \leq j \implies \downarrow k (Suc j) T[i \mapsto \tau Top]\tau = \downarrow j k (T[i \mapsto \tau Top]\tau) \]
(\text{proof})

lemma substE-substE:
\[ i \leq j \implies \Gamma[Suc j \mapsto \tau V]\tau[i \mapsto \tau U]\tau[j - i \mapsto \tau V]\tau = \Gamma[i \mapsto \tau U]\tau[j \mapsto \tau V]\tau \]
(\text{proof})

lemma substT-decE [simp]:
\[ i \leq j \implies \downarrow c k (Suc j) \Gamma[i \mapsto \tau Top]\tau = \downarrow c j k (\Gamma[i \mapsto \tau Top]\tau) \]
(\text{proof})

lemma liftE-app [simp]: \[ \uparrow c n k (\Gamma \otimes \Delta) = \uparrow c n (k + \|\Delta\|) \Gamma \otimes \uparrow c n k \Delta \]
(\text{proof})

lemma substE-app [simp]:
\[ (\Gamma \otimes \Delta)[k \mapsto \tau T]\tau\tau = \Gamma[k + \|\Delta\| \mapsto \tau T]\tau\tau \otimes \Delta[k \mapsto \tau T]\tau\tau \]
(\text{proof})

lemma substs-app [simp]:
\[ t[k \mapsto_s \tau s \otimes \tau s][\|\tau s\| \mapsto \tau s][\tau s] = t[k + \|\tau s\| \mapsto_s \tau s][k \mapsto_s \tau s] \]
(\text{proof})

theorem decE-Nil [simp]:
\[ \downarrow c n k [] = [] \]
(\text{proof})

theorem decE-Cons [simp]:
\[ \downarrow c n k (B :: \Gamma) = \text{mapB}_{\downarrow c n (k + \|\Gamma\|)} B :: \downarrow c n k \Gamma \]
(\text{proof})

theorem decE-app [simp]:
\[ \downarrow c n k (\Gamma \otimes \Delta) = \downarrow c n (k + \|\Delta\|) \Gamma \otimes \downarrow c n k \Delta \]
(\text{proof})

theorem decT-liftT [simp]:
\[ k \leq k' \implies k' + m \leq k + n \implies \downarrow \tau m k' (\uparrow \tau n k \Gamma) = \uparrow \tau (n - m) k \Gamma \]
\[\begin{align*}
\text{theorem } \text{decE-liftE } \text{[simp]}: &\quad k \leq k' \implies k' + m \leq k + n \implies \downarrow_e m \ k' \ (\uparrow_e n \ k \ \Gamma) = \uparrow_e (n - m) \ k \ \Gamma \\
\text{lemma } \text{decT-decT } \text{[simp]}: &\quad \downarrow_T n \ k \ (\downarrow_T n' (k + n) \ T) = \downarrow_T (n + n') \ k \ T \\
\text{lemma } \text{decE-decE } \text{[simp]}: &\quad \downarrow_e n \ k \ (\downarrow_e n' (k + n) \ \Gamma) = \downarrow_e (n + n') \ k \ \Gamma \\
\text{lemma } \text{decE-length } \text{[simp]}: &\quad \|\downarrow_e n \ k \ \Gamma\| = \|\Gamma\| \\
\text{lemma } \text{liftrT-assoc-None} \text{ [simp]}: &\quad (\uparrow_{\tau\tau} n \ k \ fs(l)\gamma = \bot) = (fs(l)\gamma = \bot) \\
\text{lemma } \text{liftrT-assoc-Some} \text{ [simp]}: &\quad (\uparrow_{\tau\tau} n \ k \ fs(l)\gamma = \bot) = (fs(l)\gamma = \bot) \\
\text{lemma } \text{liftr-assoc-None} \text{ [simp]}: &\quad (\uparrow_{\tau} n \ k \ fs(l)\gamma = \bot) = (fs(l)\gamma = \bot) \\
\text{lemma } \text{unique-liftrT} \text{ [simp]}: &\quad \text{unique } (\uparrow_{\tau\tau} n \ k \ fs) = \text{unique } fs \\
\text{lemma } \text{substrTT-assoc-None} \text{ [simp]}: &\quad (fs(k \mapsto \tau \ U)_{r\tau}(a)\gamma = \bot) = (fs(a)\gamma = \bot) \\
\text{lemma } \text{substrTT-assoc-Some} \text{ [simp]}: &\quad (fs(k \mapsto \tau \ U)_{r\tau}(a)\gamma = \bot) = (fs(a)\gamma = \bot) \\
\text{lemma } \text{substrT-assoc-None} \text{ [simp]}: &\quad (fs(k \mapsto \tau \ U)_{r\tau}(l)\gamma = \bot) = (fs(l)\gamma = \bot) \\
\text{lemma } \text{substrT-assoc-Some} \text{ [simp]}: &\quad (fs(k \mapsto \tau \ U)_{r\tau}(l)\gamma = \bot) = (fs(l)\gamma = \bot) \\
\text{lemma } \text{substr-assoc-None} \text{ [simp]}: &\quad (fps(k \mapsto \tau \ U)_{r\tau}(l)\gamma = \bot) = (fps(l)\gamma = \bot) \\
\text{lemma } \text{substrp-assoc-None} \text{ [simp]}: &\quad (fps(k \mapsto \tau \ U)_{r\tau}(l)\gamma = \bot) = (fps(l)\gamma = \bot) \\
\end{align*}\]
lemma unique-substrT [simp]: unique (fs[k \mapsto \tau]_\tau) = unique fs
(proof)

lemma liftrT-set: (a, T) \in set fs \implies (a, \uparrow n k T) \in set (\uparrow n k T)
(proof)

lemma liftrT-setD:
(a, T) \in set (\uparrow n k T) \implies \exists T'. (a, T') \in set (\uparrow n k T')
(proof)

lemma substrT-set: (a, T) \in set fs \implies (a, T[k \mapsto \tau]_\tau) \in set (fs[k \mapsto \tau]_\tau)
(proof)

lemma substrT-setD:
(a, T) \in set (fs[k \mapsto \tau]_\tau) \implies \exists T'. (a, T') \in set (fs[k \mapsto \tau]_\tau)
(proof)

3.3 Well-formedness

The definition of well-formedness is extended with a rule stating that a record type \text{RcdT} fs is well-formed, if for all fields \((l, T)\) contained in the list \(fs\), the type \(T\) is well-formed, and all labels \(l\) in \(fs\) are \textit{unique}.

\textbf{inductive} well-formed :: env \Rightarrow type \Rightarrow bool (- \vdash_{wf} [- [50, 50] 50])
\textbf{where}
- \textbf{wf-TVar}: \Gamma(i) = [TVarB T] \implies \Gamma \vdash_{wf} TVar i
- \textbf{wf-Top}: \Gamma \vdash_{wf} Top
- \textbf{wf-arrow}: \Gamma \vdash_{wf} T U \implies \Gamma \vdash_{wf} T \rightarrow U
- \textbf{wf-all}: \Gamma \vdash_{wf} T \Rightarrow TVarB T :: \Gamma \vdash_{wf} U \implies \Gamma \vdash_{wf} (\forall <:T, U)
- \textbf{wf-RcdT}: unique fs \implies \forall (l, T) \in set fs. \Gamma \vdash_{wf} T \Rightarrow \Gamma \vdash_{wf} \text{RcdT} fs

\textbf{inductive} well-formedE :: env \Rightarrow bool (- \vdash_{wf} [50] 50)
\textbf{and} well-formedB :: env \Rightarrow binding \Rightarrow bool (- \vdash_{wf} B [- [50, 50] 50])
\textbf{where}
- \textbf{wf-Nil}: [] \vdash_{wf}
- \textbf{wf-Cons}: \Gamma \vdash_{wf} B \implies \Gamma \vdash_{wf} B :: \Gamma \vdash_{wf}

\textbf{inductive-cases} well-formed-cases:
- \Gamma \vdash_{wf} TVar i
- \Gamma \vdash_{wf} Top
- \Gamma \vdash_{wf} T \rightarrow U
- \Gamma \vdash_{wf} (\forall <:T, U)
- \Gamma \vdash_{wf} (\text{RcdT} fTs)

\textbf{inductive-cases} well-formedE-cases:
- \textbf{B :: \Gamma \vdash_{wf}}

25
lemma wf-TVarB: \( \Gamma \vdash_{wf} T \Longrightarrow \Gamma \vdash_{wf} T \text{VarB} \text{ } T : : \Gamma \vdash_{wf} \)

(proof)

lemma wf-VarB: \( \Gamma \vdash_{wf} T \Longrightarrow \Gamma \vdash_{wf} \text{VarB} \text{ } T : : \Gamma \vdash_{wf} \)

(proof)

lemma map-is-TVarB:
\[
\text{map is-TVarB } \Gamma' = \text{map is-TVarB } \Gamma \Longrightarrow \\
\Gamma'(i) = \lceil T \text{VarB } T \rceil \Longrightarrow \exists T. \Gamma'(i) = \lceil T \text{VarB } T \rceil
\]

(proof)

lemma wf-equal-length:
\[
\text{assumes } H: \Gamma \vdash_{wf} T \\
\text{shows } \text{map is-TVarB } \Gamma' = \text{map is-TVarB } \Gamma \Longrightarrow \Gamma' \vdash_{wf} T
\]

(proof)

lemma wfE-replace:
\[
\Delta @ B : : \Gamma \vdash_{wf} B' \Longrightarrow \text{is-TVarB } B' = \text{is-TVarB } B \Longrightarrow \\
\Delta @ B' : : \Gamma \vdash_{wf}
\]

(proof)

lemma wf-weaken:
\[
\text{assumes } H: \Delta @ \Gamma \vdash_{wf} T \\
\text{shows } \uparrow_e (\text{Suc } 0) 0 \Delta @ B : : \Gamma \vdash_{wf} \uparrow_e (\text{Suc } 0) \| \Delta \| \| T 
\]

(proof)

lemma wf-weaken': \( \Gamma \vdash_{wf} T \Longrightarrow \Delta @ \Gamma \vdash_{wf} \uparrow_e (\text{Suc } i) 0 \| T \)

(proof)

lemma wfE-weaken: \( \Delta @ \Gamma \vdash_{wf} \Longrightarrow \Gamma \vdash_{wf} B \Longrightarrow \uparrow_e (\text{Suc } 0) 0 \Delta @ B : : \Gamma \vdash_{wf} \)

(proof)

lemma wf-liftB:
\[
\text{assumes } H: \Gamma \vdash_{wf} \\
\text{shows } \Gamma \langle i \rangle = \lfloor \text{VarB } T \rfloor \Longrightarrow \Gamma \vdash_{wf} \uparrow_e (\text{Suc } i) 0 \| T 
\]

(proof)

theorem wf-subst:
\[
\Delta @ B : : \Gamma \vdash_{wf} T \Longrightarrow \Gamma \vdash_{wf} U \Longrightarrow \Delta[\theta \mapsto_T U]_e @ \Gamma \vdash_{wf} T[\| \Delta \| \mapsto_T U]_{\tau} \\
\forall (l, T) \in \text{set } (rT::\text{redT}). \Delta @ B : : \Gamma \vdash_{wf} T \Longrightarrow \Gamma \vdash_{wf} U \Longrightarrow \\
\forall (l, T) \in \text{set } rT. \Delta[\theta \mapsto_T U]_e @ \Gamma \vdash_{wf} T[\| \Delta \| \mapsto_T U]_{\tau} \\
\Delta @ B : : \Gamma \vdash_{wf} \text{snd } (fT::\text{fldT}) \Longrightarrow \Gamma \vdash_{wf} U \Longrightarrow \\
\Delta[0 \mapsto_T U]_e @ \Gamma \vdash_{wf} \text{snd } fT[\| \Delta \| \mapsto_T U]_{\tau}
\]

(proof)

theorem wf-dec:
\[
\Delta @ B : : \Gamma \vdash_{wf} T \Longrightarrow \Gamma \vdash_{wf} \downarrow \| \Delta \| \| 0 \| T
\]

(proof)

theorem wfE-subst: \( \Delta @ B : : \Gamma \vdash_{wf} \Longrightarrow \Gamma \vdash_{wf} U \Longrightarrow \Delta[\theta \mapsto_T U]_e @ \Gamma \vdash_{wf} \)


3.4 Subtyping

The definition of the subtyping judgement is extended with a rule $SA-Rcd$ stating that a record type $RcdT \; fs$ is a subtype of $RcdT \; fs'$, if for all fields $(l, \; T)$ contained in $fs'$, there exists a corresponding field $(l, \; S)$ such that $S$ is a subtype of $T$. If the list $fs'$ is empty, $SA-Rcd$ can appear as a leaf in the derivation tree of the subtyping judgement. Therefore, the introduction rule needs an additional premise $\Gamma \vdash wf$ to make sure that only subtyping judgements with well-formed contexts are derivable. Moreover, since $fs$ can contain additional fields not present in $fs'$, we also have to require that the type $RcdT \; fs$ is well-formed. In order to ensure that the type $RcdT \; fs'$ is well-formed, too, we only have to require that labels in $fs'$ are unique, since, by induction on the subtyping derivation, all types contained in $fs'$ are already well-formed.

\[
\text{inductive} \quad \text{subtyping} :: \text{env} \Rightarrow \text{type} \Rightarrow \text{type} \Rightarrow \text{bool} \ (-\vdash -::<\{50, 50, 50\} 50) \\
\text{where} \\
SA-\text{Top}: \Gamma \vdash wf \implies \Gamma \vdash wf \; S \implies \Gamma \vdash S <: \text{Top} \\
| \; \text{SA-refl-TVar}: \Gamma \vdash wf \implies \Gamma \vdash wf \; \text{TVar} \; i \implies \Gamma \vdash \text{TVar} \; i <: \text{TVar} \; i \\
| \; \text{SA-trans-TVar}: \Gamma \langle i \rangle \; \Gamma \downarrow \Gamma \downarrow (\text{Suc} \; i) \; \text{0} \; U <: \Gamma \implies \Gamma \vdash \text{TVar} \; i <: \text{TVar} \; i \\
| \; \text{SA-all}: \Gamma \vdash T_1 <: S_1 \implies \Gamma \vdash T_2 <: S_2 \implies \Gamma \vdash S_1 <: S_2 \implies \Gamma \vdash T_1 <: T_2 \\
| \; \text{SA-Rcd}: \Gamma \vdash wf \implies \Gamma \vdash wf \; \text{RcdT} \; fs \implies \text{unique} \; fs' \implies \\
\forall (l, \; T) \in \text{set} \; fs'. \; \exists \; S. \; (l, \; S) \in \text{set} \; fs \land \Gamma \vdash S <: T \implies \Gamma \vdash \text{RcdT} \; fs <: \text{RcdT} \; fs'
\]

\text{lemma} \; \text{wf-subtype-env}:
\quad assumes \; PQ: \Gamma \vdash P <: Q \\
\quad shows \; \Gamma \vdash wf \langle \text{proof} \rangle

\text{lemma} \; \text{wf-subtype}:
\quad assumes \; PQ: \Gamma \vdash P <: Q \\
\quad shows \; \Gamma \vdash wf \; P \land \Gamma \vdash wf \; Q \langle \text{proof} \rangle

\text{lemma} \; \text{wf-subtypeE}:
\quad assumes \; H: \Gamma \vdash T <: U \\
\quad and \; H': \Gamma \vdash wf \implies \Gamma \vdash wf \; T \implies \Gamma \vdash wf \; U \implies P \\
\quad shows \; P \\
\langle \text{proof} \rangle

\text{lemma} \; \text{subtype-refl}: \quad \Lambda.1 \\
\quad \Gamma \vdash wf \implies \Gamma \vdash wf \; T \implies \Gamma \vdash T <: T \\
\quad \Gamma \vdash wf \implies \forall (l::\text{name}, \; T) \in \text{set} \; fTs. \; \Gamma \vdash wf \; T \implies \Gamma \vdash T <: T \\
\quad \Gamma \vdash wf \implies \Gamma \vdash wf \; \text{snd} \; (\text{ft}::\text{fldT}) \implies \Gamma \vdash \text{snd} \; \text{ft} <: \text{snd} \; \text{ft}
\begin{proof}

**Lemma subtype-weaken:**

Assumes $H: \Delta \at \Gamma \vdash P \triangleq Q$ and $wf: \Gamma \vdash w_f B$

Shows $\downarrow e \ 0 \ \Delta \at B :: \Gamma \vdash \uparrow \tau \ \| \Delta \| \ P \triangleq \uparrow \tau \ \| \Delta \| \ Q$

(\proof)

**Lemma subtype-weaken':** — A.2

$\Gamma \vdash P \triangleq Q \implies \Delta \at \Gamma \vdash wf \triangleq \Delta \at \Gamma \vdash \uparrow \tau \ \| \Delta \| \ 0 \ P \triangleq \uparrow \tau \ \| \Delta \| \ 0 \ Q$

(\proof)

**Lemma fieldT-size [simp]:**

$(a, T) \in set fs \implies size T < Suc (size-list (size-prod (\lambda x. 0) size) fs)$

(\proof)

**Lemma subtype-trans:** — A.3

$\Gamma \vdash S \triangleq Q \implies \Gamma \vdash Q \triangleq T \implies \Gamma \vdash S \triangleq T$

$\Delta \at TVarB Q :: \Gamma \vdash M \triangleq N \implies \Gamma \vdash P \triangleq Q \implies \Delta \at TVarB P :: \Gamma \vdash M \triangleq N$

(\proof)

**Lemma substT-subtype:** — A.10

Assumes $H: \Delta \at TVarB Q :: \Gamma \vdash S \triangleq T$

Shows $\downarrow e \ 0 \ \Delta \at \Gamma \vdash \downarrow \tau \ 1 \ \| \Delta \| \ 0 \ P \triangleq \downarrow \tau \ 1 \ \| \Delta \| \ 0 \ Q$

(\proof)

**Lemma subst-subtype:**

Assumes $H: \Delta \at VarB V :: \Gamma \vdash T \triangleq U$

Shows $\downarrow e \ 0 \ \Delta \at \Gamma \vdash \downarrow \tau \ 1 \ \| \Delta \| \ T \triangleq \downarrow \tau \ 1 \ \| \Delta \| \ U$

(\proof)

### 3.5 Typing

In the formalization of the type checking rule for the LET binder, we use an additional judgement $\vdash p : T \Rightarrow \Delta$ for checking whether a given pattern $p$ is compatible with the type $T$ of an object that is to be matched against this pattern. The judgement will be defined simultaneously with a judgement $\vdash ps [:] Ts \Rightarrow \Delta$ for type checking field patterns. Apart from checking the type, the judgement also returns a list of bindings $\Delta$, which can be thought of as a “flattened” list of types of the variables occurring in the pattern. Since typing environments are extended “to the left”, the bindings in $\Delta$ appear in reverse order.

**Inductive**

\begin{align*}
ptyping :: & pat \Rightarrow \text{type} \Rightarrow \text{env} \Rightarrow \text{bool} \ (\vdash - : \cdot \Rightarrow - \ [50, 50, 50]) \\
\text{and} \ ptypings :: & rpatt \Rightarrow \text{rdt} \Rightarrow \text{env} \Rightarrow \text{bool} \ (\vdash - [:] \Rightarrow - \ [50, 50, 50]) \\
\text{where} \ \\
P-\text{Var}: & \vdash P\text{Var} T : T \Rightarrow [\text{VarB} \ T]
\end{align*}
lemmas ptyping-induct = ptyping-ptytypings.inducts(1)
[of - - - \lambda x y z. True, simplified True-simps, consumes 1, case-names P-Var P-Red]

lemmas ptytings-induct = ptyping-ptytings.inducts(2)
[of - - - \lambda x y z. True, simplified True-simps, consumes 1, case-names P-Nil P-Cons]

lemmas typing-induct = typing-typings.inducts(1)
[of - - - \lambda x y z. True, simplified True-simps, consumes 1, case-names T-Var T-Abs T-App T-TAbs T-TApp T-Sub T-Let T-Rcd T-Proj]

lemmas typings-induct = typing-typings.inducts(2)
[of - - - \lambda x y z. True, simplified True-simps, consumes 1, case-names T-Nil T-Cons]

lemma narrow-type: — \Lambda.7
\[ \Delta @ TVarB Q :: \Gamma \vdash t : T \implies \Gamma \vdash P <: Q \implies \Delta @ TVarB P :: \Gamma \vdash t : T \]
\[ \Delta @ TVarB Q :: \Gamma \vdash ts [\cdot] Ts \implies \Gamma \vdash P <: Q \implies \Delta @ TVarB P :: \Gamma \vdash ts [\cdot] Ts \]
(proof)

lemma typings-setD:
assumes H: \Gamma \vdash fs [\cdot] fTs
shows \((l, T) \in \text{set} fTs \implies \exists t. fs(l) = [t] \land \Gamma \vdash t : T \)
(proof)

lemma subtype-refl':
assumes t: \Gamma \vdash t : T
shows \(\Gamma \vdash T <: T \)
(proof)

lemma Abs-type: — A.13(1)
assumes H: \Gamma \vdash (\lambda S. s) : T
shows \(\Gamma \vdash T <: (\forall S'. \Gamma \vdash U <: S \implies \text{Var} B S :: \Gamma \vdash s : S' \implies \Gamma \vdash \downarrow 1 \ 0 S' <: U' \implies P) \implies P \)
(proof)

lemma Abs-type':
assumes H: \Gamma \vdash (\lambda S. s) : \Sigma \to US
and R: \(\forall S'. \Gamma \vdash U <: S \implies \text{Var} B S :: \Gamma \vdash s : S' \implies \Gamma \vdash \downarrow 1 \ 0 S' <: U' \implies P \)
shows P (proof)

lemma TAbs-type: — A.13(2)
assumes H: \Gamma \vdash (\lambda< U. s) : T
shows \(\Gamma \vdash T <: (\forall U'. \Gamma \vdash s : (U < U'. U') \implies \text{Var} B U :: \Gamma \vdash s : S' \implies TVarB U :: \Gamma \vdash S' <: U' \implies P) \implies P \)
(proof)

lemma TAbs-type':
assumes H: \Gamma \vdash (\lambda< U. s) : (\forall U'. U')
and R: \(\forall S'. \Gamma \vdash U <: S \implies \text{Var} B U :: \Gamma \vdash s : S' \implies TVarB U :: \Gamma \vdash S' <: U' \implies P \)
shows P (proof)
In the proof of the preservation theorem, the following elimination rule for typing judgements on record types will be useful:

**Lemma Red-type1**: \[ \begin{array}{l}
\text{Assumes } H: \Gamma \vdash t : T \\
\text{Shows } t = \text{Red } fs \implies \Gamma \vdash T <: \text{Red } fTs \implies \\
\forall (l, U) \in \text{set } fTs. \exists u. fs(l) \equiv [u] \land \Gamma \vdash u : U
\end{array} \]

**Proof**

**Lemma Red-type1’**: \[ \begin{array}{l}
\text{Assumes } H: \Gamma \vdash \text{Red } fs : \text{Red } fTs \\
\text{Shows } \forall (l, U) \in \text{set } fTs. \exists u. fs(l) \equiv [u] \land \Gamma \vdash u : U
\end{array} \]

Intuitively, this means that for a record \( \text{Red } fs \) of type \( \text{Red } fTs \), each field with name \( l \) associated with a type \( U \) in \( fTs \) must correspond to a field in \( fs \) with value \( u \), where \( u \) has type \( U \). Thanks to the subsumption rule \( T-\text{Sub} \), the typing judgement for terms is not sensitive to the order of record fields. For example,

\[ \Gamma \vdash \text{Red } [(l_1, t_1), (l_2, t_2), (l_3, t_3)] : \text{Red } T [(l_2, T_2), (l_1, T_1)] \]

provided that \( \Gamma \vdash t_i : T_i \). Note however that this does not imply

\[ \Gamma \vdash [(l_1, t_1), (l_2, t_2), (l_3, t_3)] :: [(l_2, T_2), (l_1, T_1)] \]

In order for this statement to hold, we need to remove the field \( l_3 \) and exchange the order of the fields \( l_1 \) and \( l_2 \). This gives rise to the following variant of the above elimination rule:

**Lemma Red-type2**: \[ \begin{array}{l}
\Gamma \vdash \text{Red } fs : T \implies \Gamma \vdash T <: \text{Red } fTs \implies \\
\Gamma \vdash \text{map} (\lambda (l, T). (l, \text{the } (fs(l)))) fTs :: fTs
\end{array} \]

**Proof**

**Lemma Red-type2’**: \[ \begin{array}{l}
\text{Assumes } H: \Gamma \vdash \text{Red } fs : \text{Red } fTs \\
\text{Shows } \Gamma \vdash \text{map} (\lambda (l, T). (l, \text{the } (fs(l)))) fTs :: fTs
\end{array} \]

**Lemma T-eq**: \[ \begin{array}{l}
\Gamma \vdash t : T \implies T = T' \implies \Gamma \vdash t : T' \ (\text{proof})
\end{array} \]

**Lemma ptyping-length**: \[ \begin{array}{l}
\vdash p : T \Rightarrow \Delta \implies \|p\|_p = \|\Delta\|
\end{array} \]

**Proof**

**Lemma lift-ptyping**: \[ \begin{array}{l}
\vdash p : T \Rightarrow \Delta \implies \uparrow_p n \ k \ p : \uparrow r_n k \ T \Rightarrow \uparrow c n k \ \Delta
\end{array} \]

\[ \vdash \text{fps} :: fTs \Rightarrow \Delta \implies \uparrow r_p n \ k \ \text{fps} :: \uparrow r r_n k \ fTs \Rightarrow \uparrow c n k \ \Delta \]
proof}

lemma type-weaken:
\[ \Delta @ \Gamma \vdash t : T \Longrightarrow \Gamma \vdash \text{wf}_B B \Longrightarrow \]
\[ \uparrow_e I 0 \Delta @ B :: \Gamma \vdash \uparrow I \| \Delta \| I \| T \| \]
\[ \Delta @ \Gamma \vdash \text{fs} [:] fTs \Longrightarrow \Gamma \vdash \text{wf}_B B \Longrightarrow \]
\[ \uparrow_e I 0 \Delta @ B :: \Gamma \vdash \uparrow \| \Delta \| fTs \]
\( \langle \text{proof} \rangle \)

lemma type-weaken\': — A.5(6)
\[ \Delta @ \Gamma \vdash t : T \Longrightarrow \Delta @ \Gamma \vdash \text{wf} \Longrightarrow \Delta @ \Gamma \vdash \uparrow \| \Delta \| 0 t : \uparrow \| \Delta \| 0 T \]
\( \langle \text{proof} \rangle \)

The substitution lemmas are now proved by mutual induction on the derivations of the typing derivations for terms and lists of fields.

lemma subst-ptyping:
\[ \vdash p : T \Longrightarrow \Delta \vdash \text{fstype} \vdash \Delta @ \Gamma \vdash \uparrow \| \Delta \| 0 t : \uparrow \| \Delta \| 0 T \]
\( \langle \text{proof} \rangle \)

theorem subst-type: — A.8
\[ \Delta @ \text{Var}_B U :: \Gamma \vdash t : T \Longrightarrow \Gamma \vdash u : U \Longrightarrow \]
\[ \downarrow_e I 0 \Delta @ \Gamma \vdash \uparrow \| \Delta \| \vdash u : \uparrow \| \Delta \| \]
\[ \Delta @ \text{Var}_B U :: \Gamma \vdash \text{fstype} :: \Gamma \vdash u : U \Longrightarrow \]
\[ \downarrow_e I 0 \Delta @ \Gamma \vdash \text{fstype} :: \Gamma \vdash u : U \]
\( \langle \text{proof} \rangle \)

theorem substT-type: — A.11
\[ \Delta @ \text{Var}_B Q :: \Gamma \vdash t : T \Longrightarrow \Gamma \vdash P : Q \Longrightarrow \]
\[ \Delta[\theta \mapsto \tau]_e @ \Gamma \vdash t[\| \Delta \| \mapsto \tau] \vdash P[\| \Delta \| \mapsto \tau]_e \]
\[ \Delta @ \text{Var}_B Q :: \Gamma \vdash \text{fstype} :: \Gamma \vdash P : Q \Longrightarrow \]
\[ \Delta[\theta \mapsto \tau]_e @ \Gamma \vdash \text{fstype} :: \Gamma \vdash P[\| \Delta \| \mapsto \tau] \vdash P[\| \Delta \| \mapsto \tau]_e \]
\( \langle \text{proof} \rangle \)

3.6 Evaluation

The definition of canonical values is extended with a clause saying that a record Red fs is a canonical value if all fields contain canonical values:

inductive-set
\[ \text{value} :: \text{trm} \set \]
where
\[ \text{Abs}: (\lambda : T. \ t) \in \text{value} \]
| \[ T\set : (\lambda : T. \ t) \in \text{value} \]
\[ \text{Red}: \forall (t, \ t) \in \text{set} \set. \ t \in \text{value} \Longrightarrow \text{Red} \set \in \text{value} \]

In order to formalize the evaluation rule for LET, we introduce another relation \( \vdash p \rightarrow t \rightarrow ts \) expressing that a pattern \( p \) matches a term \( t \). The relation also yields a list of terms \( ts \) corresponding to the variables in the
pattern. The relation is defined simultaneously with another relation \( \vdash fps [\triangleright] fs \Rightarrow ts \) for matching a list of field patterns \( fps \) against a list of fields \( fs \):

**inductive**

\[
\text{match :: pat} \Rightarrow \text{trm} \Rightarrow \text{trm list} \Rightarrow \text{bool} \quad (\vdash - \triangleright - \Rightarrow - [50, 50, 50])
\]

**and** matches :: rpat \Rightarrow \text{rcd} \Rightarrow \text{trm list} \Rightarrow \text{bool} \quad (\vdash [\triangleright] - \Rightarrow - [50, 50, 50])

**where**

\[
\begin{align*}
\text{M-PVar:} & \vdash \text{PVar } T \triangleright t \Rightarrow [t] \\
\text{M-Rcd:} & \vdash fps [\triangleright] fs \Rightarrow ts \Rightarrow \vdash \text{PRcd} \triangleright \text{Rcd } fs \Rightarrow ts \\
\text{M-Nil:} & \vdash [] [\triangleright] fs \Rightarrow [] \\
\text{M-Cons:} & \vdash fs(l)\gamma = [t] \Rightarrow \vdash p \triangleright t \Rightarrow ts \Rightarrow \vdash fps [\triangleright] fs \Rightarrow us \Rightarrow \\
& \vdash (l, p) :: fps [\triangleright] fs \Rightarrow ts \& us
\end{align*}
\]

The rules of the evaluation relation for the calculus with records are as follows:

**inductive**

\[
\text{eval :: trm} \Rightarrow \text{trm} \Rightarrow \text{bool} \quad (\text{infixl} \mapsto 50)
\]

**and** evals :: recd \Rightarrow recd \Rightarrow bool \quad (\text{infixl} \mapsto 50)

**where**

\[
\begin{align*}
\text{E-Abs:} & \quad v_2 \in \text{value} \Rightarrow (\lambda : T_{11}, t_{12}) \cdot v_2 \Rightarrow t_{12}[0 \mapsto v_2] \\
\text{E-TAbs:} & \quad (\lambda <: T_{11}, t_{12} \cdot \tau \ T_2 \mapsto t_{12}[0 \mapsto \tau \ T_2] \\
\text{E-App:} & \quad t \mapsto t' \Rightarrow t \cdot u \mapsto t' \cdot u \\
\text{E-App2:} & \quad v \in \text{value} \Rightarrow \vdash t \mapsto t' \Rightarrow v \cdot t \mapsto v \cdot t' \\
\text{E-TApp:} & \quad t \mapsto t' \Rightarrow t \cdot \tau \ T \mapsto t' \cdot \tau \ T \\
\text{E-LetV:} & \quad v \in \text{value} \Rightarrow \vdash \triangleright p \triangleright v \Rightarrow ts \Rightarrow (\text{LET } p = v \text{ IN } t) \mapsto t[0 \mapsto_s ts] \\
\text{E-ProjRcd:} & \quad fs(l)\gamma = [v] \Rightarrow v \in \text{value} \Rightarrow \text{Rcd } fs.l \mapsto v \\
\text{E-Proj:} & \quad t \mapsto t' \Rightarrow t.l \mapsto t'.l \\
\text{E-Rcd:} & \quad fs [\mapsto] fs' \Rightarrow \text{Rcd } fs \mapsto \text{Rcd } fs' \\
\text{E-Let:} & \quad t \mapsto t' \Rightarrow (\text{LET } p = t \text{ IN } u) \mapsto (\text{LET } p = t' \text{ IN } u) \\
\text{E-hd:} & \quad t \mapsto t' \Rightarrow (l, t) :: fs [\mapsto] (l, t') :: fs \\
\text{E-tl:} & \quad v \in \text{value} \Rightarrow fs [\mapsto] fs' \mapsto (l, v) :: fs [\mapsto] (l, v) :: fs'
\end{align*}
\]

The relation \( t \mapsto t' \) is defined simultaneously with a relation \( fs [\mapsto] fs' \) for evaluating record fields. The “immediate” reductions, namely pattern matching and projection, are described by the rules E-LetV and E-ProjRcd, respectively, whereas E-Proj, E-Rcd, E-Let, E-hd and E-tl are congruence rules.

**lemmas** matches-induct = match-matches.inducts(2)

[of - - \lambda x y z. True, simplified True-simps, consumes 1, case-names M-Nil M-Cons]

**lemmas** evals-induct = eval-evals.inducts(2)

[of - \lambda x y. True, simplified True-simps, consumes 1, case-names E-hd E-tl]

**lemma** matches-mono:

assumes \( H: \vdash fps [\triangleright] fs \Rightarrow ts \)

shows \( fps(l)\gamma = \bot \Rightarrow \vdash fps [\triangleright] (l, t) :: fs \Rightarrow ts \)

(proof)

33
lemma matches-eq:
assumes $H: \vdash fps \rightarrow fps \Rightarrow ts$
shows $\forall (l, p) \in \text{set} \ fps. \ fs(l)\gamma = fs'(l)\gamma \Longrightarrow \vdash fps \rightarrow fps' \Rightarrow ts$
(\text{proof})

lemma reorder-eq:
assumes $H: \vdash fps \rightarrow \Delta$
shows $\forall (l, U) \in \text{set} \ fps. \ \exists u. \ fs(l)\gamma = [u] \Longrightarrow \forall (l, p) \in \text{set} \ fps. \ fs(l)\gamma = (\text{map} (\lambda(l, T). \ (l, \text{the} \ (fs(l)\gamma))) \ fps\rightarrow ts \gamma)\Rightarrow$
(\text{proof})

lemma matches-reorder:
$\vdash fps \rightarrow \Delta \Rightarrow \forall (l, U) \in \text{set} \ fps. \ \exists u. \ fs(l)\gamma = [u] \Longrightarrow \vdash fps \rightarrow ts \Rightarrow \vdash fps \rightarrow ts$
(\text{proof})

lemma matches-reorder':
$\vdash fps \rightarrow \Delta \Rightarrow \forall (l, U) \in \text{set} \ fps. \ \exists u. \ fs(l)\gamma = [u] \Longrightarrow \vdash fps \rightarrow ts \Rightarrow \vdash fps \rightarrow ts$
(\text{proof})

theorem matches-tl:
assumes $H: \vdash fps \rightarrow (l, t) :: fs \Rightarrow ts$
shows $fps(l)\gamma = \bot \Longrightarrow \vdash fps \rightarrow ts$
(\text{proof})

theorem match-length:
$\vdash p \triangleright t \Rightarrow ts \Longrightarrow \vdash p : T \Rightarrow \Delta \Longrightarrow ||ts|| = ||\Delta||$
$\vdash fps \rightarrow ft \Rightarrow ts \Longrightarrow \vdash fps \rightarrow fps \rightarrow \Delta \Longrightarrow ||ts|| = ||\Delta||$
(\text{proof})

In the proof of the preservation theorem for the calculus with records, we need the following lemma relating the matching and typing judgements for patterns, which means that well-typed matching preserves typing. Although this property will only be used for $\Gamma_1 = []$ later, the statement must be proved in a more general form in order for the induction to go through.

theorem match-type: $- A.17$
$\vdash p : T_1 \Rightarrow \Delta \Longrightarrow \Gamma_2 \vdash t_1 : T_1 \Longrightarrow \Gamma_1 \odot \Delta \odot \Gamma_2 \vdash t_2 : T_2 \Longrightarrow \vdash p \triangleright t_1 \Rightarrow ts \Longrightarrow \downarrow ||\Delta|| \ 0 \ \Gamma_1 \odot \Gamma_2 \vdash t_2 ||\Gamma_1|| \rightarrow_s ts \ : \downarrow ||\Delta|| \ ||\Gamma_1|| \ T_2$
$\vdash fps \rightarrow fps \rightarrow \Delta \Longrightarrow \Gamma_2 \vdash fs \rightarrow ft \rightarrow \Delta \Longrightarrow \Gamma_1 \odot \Delta \odot \Gamma_2 \vdash t_2 : T_2 \Longrightarrow \vdash fps \rightarrow ft \rightarrow ts \Longrightarrow \downarrow ||\Delta|| \ 0 \ \Gamma_1 \odot \Gamma_2 \vdash t_2 ||\Gamma_1|| \rightarrow_s ts \ : \downarrow ||\Delta|| \ ||\Gamma_1|| \ T_2$
(\text{proof})

lemma evals-labels [simp]:
assumes $H: fs \rightarrow fps'$
shows $\forall (l)\gamma = \bot \Longrightarrow (fs'(l)\gamma = \bot)$ (\text{proof})
\textbf{Theorem} preservation: — A.20
\[ \Gamma \vdash t : T \implies t \rightsquigarrow t' \implies \Gamma \vdash t' : T \]
\[ \Gamma \vdash fs [\cdot] fTs \implies fs [\rightsquigarrow] fs' \implies \Gamma \vdash fs' [\cdot] fTs \]
(proof)

\textbf{Lemma} Fun-canonical: — A.14(1)
\textbf{Assumes} \([\cdot] \vdash v : T_1 \rightarrow T_2\)
\textbf{Shows} \(v \in \text{value} \implies \exists S. v = (\lambda S. t)\)
(proof)

\textbf{Lemma} TyAll-canonical: — A.14(3)
\textbf{Assumes} \([\cdot] \vdash v : (\forall \cdot : T_1. T_2)\)
\textbf{Shows} \(v \in \text{value} \implies \exists t S. v = (\lambda \cdot : S. t)\)
(proof)

Like in the case of the simple calculus, we also need a canonical values theorem for record types:

\textbf{Lemma} RedT-canonical: — A.14(2)
\textbf{Assumes} \([\cdot] \vdash v : \text{RedT} fTs\)
\textbf{Shows} \(v \in \text{value} \implies \exists fs. v = \text{Red} fs \land (\forall (l, t) \in \text{set fs}. t \in \text{value})\)
(proof)

\textbf{Theorem} reorder-prop:
\[ \forall (l, t) \in \text{set fs.} P t \implies \forall (l, U) \in \text{set fTs.} \exists u. fs(l)_7 = [u] \implies \]
\[ \forall (l, t) \in \text{set (map} \ (\lambda (l, T). (l, \text{the} (fs(l)_7))) fTs). P t \]
(proof)

Another central property needed in the proof of the progress theorem is that well-typed matching is defined. This means that if the pattern \(p\) is compatible with the type \(T\) of the closed term \(t\) that it has to match, then it is always possible to extract a list of terms \(ts\) corresponding to the variables in \(p\). Interestingly, this important property is missing in the description of the PoplMark Challenge [1].

\textbf{Theorem} ptyping-match:
\[ \vdash p : T \Rightarrow \Delta \implies [\cdot] \vdash t : T \implies t \in \text{value} \implies \]
\[ \exists ts. \vdash p \triangleright t \Rightarrow ts \]
\[ \vdash fps [\cdot] fTs \Rightarrow \Delta \implies [\cdot] \vdash fs [\cdot] fTs \implies \]
\[ \forall (l, t) \in \text{set fs.} t \in \text{value} \implies \exists us. \vdash fps [\triangleright] fs \Rightarrow us \]
(proof)

\textbf{Theorem} progress: — A.16
\[ [\cdot] \vdash t : T \Rightarrow t \in \text{value} \lor (\exists t', t \rightsquigarrow t') \]
\[ [\cdot] \vdash fs [\cdot] fTs \Rightarrow (\forall (l, t) \in \text{set fs.} t \in \text{value}) \lor (\exists fs'. fs [\rightsquigarrow] fs') \]
(proof)
4 Evaluation contexts

In this section, we present a different way of formalizing the evaluation relation. Rather than using additional congruence rules, we first formalize a set of evaluation contexts, describing the locations in a term where reductions can occur. We have chosen a higher-order formalization of evaluation contexts as functions from terms to terms. We define simultaneously a set of evaluation contexts for records represented as functions from terms to lists of fields.

\[ \text{inductive-set} \]
\[ \text{ctx} :: (\text{trm} \Rightarrow \text{trm}) \text{ set} \]
\[ \text{and} \quad \text{rctxt} :: (\text{trm} \Rightarrow \text{rcd}) \text{ set} \]

\[ \text{where} \]
\[ C-\text{Hole}: (\lambda t. t) \in \text{ctx} \]
\[ C-\text{App1}: E \in \text{ctx} \Rightarrow (\lambda t. E t \cdot u) \in \text{ctx} \]
\[ C-\text{App2}: v \in \text{value} \Rightarrow E \in \text{ctx} \Rightarrow (\lambda t. E v \cdot E t) \in \text{ctx} \]
\[ C-\text{TApp}: E \in \text{ctx} \Rightarrow (\lambda t. E t \cdot \tau) T \in \text{ctx} \]
\[ C-\text{Proj}: E \in \text{ctx} \Rightarrow (\lambda t. E t \cdot \ell) \in \text{ctx} \]
\[ C-\text{Red}: E \in \text{rctxt} \Rightarrow (\lambda t. \text{Red} (E t)) \in \text{ctx} \]
\[ C-\text{Let}: E \in \text{ctx} \Rightarrow (\lambda t. \text{LET} p = E t \text{ IN} u) \in \text{ctx} \]
\[ C-\text{hd}: E \in \text{ctx} \Rightarrow (\lambda t. (\ell, E t) :: fs) \in \text{rctxt} \]
\[ C-\text{tl}: v \in \text{value} \Rightarrow E \in \text{rctxt} \Rightarrow (\lambda t. (\ell, v) :: E t) \in \text{rctxt} \]

\[ \text{lemmas} \quad \text{rctxt}\text{-induct} = \text{ctx}\text{-rctxt}\text{.inducts}(2) \]
\[ \text{[of - \text{x. True, simplified True-simps, consumes 1, case-names C-hd C-tl]} \]

\[ \text{lemma} \quad \text{rctxt-labels}: \]
\[ \text{assumes} \quad H: E \in \text{rctxt} \]
\[ \text{shows} \quad E t(\ell) = \top \Rightarrow E t'(\ell) = \top \langle \text{proof} \rangle \]

The evaluation relation \( t \rightarrow_c t' \) is now characterized by the rule \( E-\text{Ctx} \), which allows reductions in arbitrary contexts, as well as the rules \( E-\text{Abs} \), \( E-\text{TAbs} \), \( E-\text{LetV} \), and \( E-\text{ProjRcd} \) describing the “immediate” reductions, which have already been presented in §2.6 and §3.6.

\[ \text{inductive} \]
\[ \text{eval} :: \text{trm} \Rightarrow \text{trm} \Rightarrow \text{bool} \quad (\text{infixl} \rightarrow_c 50) \]

\[ \text{where} \]
\[ E-\text{Ctx}: t \rightarrow_c t' \Rightarrow E \in \text{ctx} \Rightarrow E t \rightarrow_c E t' \]
\[ E-\text{Abs}: v_2 \in \text{value} \Rightarrow (\lambda t_1. t_{12}) \cdot v_2 \rightarrow_c t_{12}[0 \mapsto v_2] \]
\[ E-\text{TAbs}: (\lambda <T_1. t_{12}) \cdot \tau T_2 \rightarrow_c t_{12}[0 \mapsto \tau T_2] \]
\[ E-\text{LetV}: v \in \text{value} \Rightarrow \Gamma p \triangleright v \Rightarrow ts \Rightarrow (\text{LET} p = v \text{ IN} t) \rightarrow_c t[0 \mapsto_s ts] \]
\[ E-\text{ProjRcd}: fs(\ell) = [v] \Rightarrow v \in \text{value} \Rightarrow \text{Red} fs..l \rightarrow_c v \]

In the proof of the preservation theorem, the case corresponding to the rule \( E-\text{Ctx} \) requires a lemma stating that replacing a term \( t \) in a well-typed term of the form \( E t \), where \( E \) is a context, by a term \( t' \) of the same type does not change the type of the resulting term \( E t' \). The proof is by mutual induction.
on the typing derivations for terms and records.

**lemma context-typing:** — A.18
\[ \Gamma \vdash \text{typing} \quad \Gamma \vdash u : T \Rightarrow E \in \text{ctxt} \Rightarrow u = E \cdot t \Rightarrow \]
\[ (\land T_0. \Gamma \vdash t : T_0 \Rightarrow \Gamma \vdash t' : T_0) \Rightarrow \Gamma \vdash E \cdot t \Rightarrow \]
\[ \Gamma \vdash \text{fs} : T \Rightarrow E_\tau \in \text{rctxt} \Rightarrow \text{fs} = E_\tau \cdot t \Rightarrow \]
\[ (\land T_0. \Gamma \vdash t : T_0 \Rightarrow \Gamma \vdash t' : T_0) \Rightarrow \Gamma \vdash E_\tau \cdot t' \]
\[ \langle \text{proof} \rangle \]

The fact that immediate reduction preserves the types of terms is proved in several parts. The proof of each statement is by induction on the typing derivation.

**theorem Abs-preservation:** — A.19(1)
assumes \( H: \Gamma \vdash (\lambda : T_{11}. t_{12}) \cdot t_2 : T \)
shows \( \Gamma \vdash t_{12}[\vartheta \mapsto t_2] : T \)
\[ \langle \text{proof} \rangle \]

**theorem TAbs-preservation:** — A.19(2)
assumes \( H: \Gamma \vdash (\lambda< : T_{11}. t_{12}) \cdot \tau_2 : T \)
shows \( \Gamma \vdash t_{12}[\vartheta \mapsto \tau_2] : T \)
\[ \langle \text{proof} \rangle \]

**theorem Let-preservation:** — A.19(3)
assumes \( H: \Gamma \vdash \text{LET p = t}_1 \text{ IN t}_2 : T \)
shows \( \vdash p \triangleright t_1 \Rightarrow ts \Rightarrow \Gamma \vdash t_2[\vartheta \mapsto \tau_2] : T \)
\[ \langle \text{proof} \rangle \]

**theorem Proj-preservation:** — A.19(4)
assumes \( H: \Gamma \vdash \text{Red fs..l} : T \)
shows \( \text{fs}(l)_? = [v] \Rightarrow \Gamma \vdash v : T \)
\[ \langle \text{proof} \rangle \]

**theorem preservation:** — A.20
assumes \( H: t \mapsto c t' \)
shows \( \Gamma \vdash t : T \Rightarrow \Gamma \vdash t' : T \)
\[ \langle \text{proof} \rangle \]

For the proof of the progress theorem, we need a lemma stating that each well-typed, closed term \( t \) is either a canonical value, or can be decomposed into an evaluation context \( E \) and a term \( t_0 \) such that \( t_0 \) is a redex. The proof of this result, which is called the decomposition lemma, is again by induction on the typing derivation. A similar property is also needed for records.

**theorem context-decomp:** — A.15
\[ [] \vdash t : T \Rightarrow \]
\[ t \in \text{value} \lor (\exists E t_0 t_0'. E \in \text{ctxt} \land t = E t_0 \land t_0 \mapsto c t_0') \]
\[ [] \vdash \text{fs} : T \Rightarrow \]
\[ (\forall (t, t) \in \text{set fs}, t \in \text{value}) \lor (\exists E t_0 t_0'. E \in \text{rctxt} \land \text{fs} = E t_0 \land t_0 \mapsto c t_0') \]
\[ \langle \text{proof} \rangle \]

37
\textbf{theorem} progress: — A.16
\begin{itemize}
  \item assumes \( H: \emptyset \vdash t : T \)
  \item shows \( t \in \text{value} \lor (\exists t'. t \leadsto_c t') \)
\end{itemize}
\langle proof \rangle

Finally, we prove that the two definitions of the evaluation relation are equivalent. The proof that \( t \leadsto_c t' \) implies \( t \leadsto t' \) requires a lemma stating that \( \leadsto \) is compatible with evaluation contexts.

\textbf{lemma} ctxt-imp-eval:
\begin{align*}
  E \in \text{ctxt} \Rightarrow t \leadsto t' \Rightarrow E t \leadsto E t' \\
  E_r \in \text{rctxt} \Rightarrow t \leadsto t' \Rightarrow E_r t \leadsto E_r t'
\end{align*}
\langle proof \rangle

\textbf{lemma} eval-evalc-eq: \((t \leadsto t') = (t \leadsto_c t')\)
\langle proof \rangle

\section{Executing the specification}

An important criterion that a solution to the \textsc{PoplMark} Challenge should fulfill is the possibility to \textit{animate} the specification. For example, it should be possible to apply the reduction relation for the calculus to example terms. Since the reduction relations are defined inductively, they can be interpreted as a logic program in the style of \textsc{Prolog}. The definition of the single-step evaluation relation presented in \S2.6 and \S3.6 is directly executable.

In order to compute the normal form of a term using the one-step evaluation relation \( \leadsto \), we introduce the inductive predicate \( t \Downarrow u \), denoting that \( u \) is a normal form of \( t \).

\textbf{inductive} norm :: \( \text{trm} \Rightarrow \text{trm} \Rightarrow \text{bool} \) (\textsf{infixl} \( \Downarrow \) 50)
\begin{align*}
  \text{where} & \\
  t \in \text{value} \Rightarrow t \Downarrow t \\
  \mid t \leadsto s \Rightarrow s \Downarrow u \Rightarrow t \Downarrow u
\end{align*}

\textbf{definition} normal-forms where
\begin{align*}
  \text{normal-forms} t \equiv \{ u. t \Downarrow u \}
\end{align*}

\textbf{lemma} [code-pred-intro Red-Nil]: \text{valuep} (\text{Red} [])
\langle proof \rangle

\textbf{lemma} [code-pred-intro Red-Cons]: \text{valuep} t \implies \text{valuep} (\text{Red} fs) \implies \text{valuep} (\text{Red} ((l, t) \# fs))
\langle proof \rangle

\textbf{lemmas} \text{valuep}.\text{intros}(1)[\text{code-pred-intro Abs}] \text{valuep}.\text{intros}(2)[\text{code-pred-intro TAbs}]

\textbf{code-pred} (\text{modes}: i => bool) \text{valuep}
\begin{proof}

\textbf{thm} valuep.equation

\textbf{code-pred} (modes: \(i \Rightarrow i \Rightarrow \text{bool}, i \Rightarrow o \Rightarrow \text{bool as normalize}\) \text{ norm} \ \text{proof})

\textbf{thm} norm.equation

\textbf{lemma} [code]:

normal-forms = set-of-pred o normalize \ \text{proof}

\textbf{lemma} [code-unfold]: \(x \in \text{value} \leftrightarrow \text{valuep} \ x\) \ \text{proof}

\textbf{definition}

\text{natT} :: \text{type} \ \text{where}

\text{natT} \equiv \forall \ <\ : \text{Top}. (\forall \ <\ : \text{TVar} \ 0. (\forall \ <\ : \text{TVar} \ 1. (\text{TVar} \ 2 \rightarrow \text{TVar} \ 1) \rightarrow \text{TVar} \ 0 \rightarrow \text{TVar} \ 1)))

\textbf{definition}

\text{fact2} :: \text{trm} \ \text{where}

\text{fact2} \equiv

\text{LET} \ P\text{Var} \ \text{natT} =

(\lambda <\ : \text{Top}. \lambda <\ : \text{TVar} \ 0. \lambda <\ : \text{TVar} \ 1. \lambda : \text{TVar} \ 2 \rightarrow \text{TVar} \ 1. \lambda : \text{TVar} \ 1. \ \text{Var} \ 1 \cdot \text{Var} \ 0)

\text{LET} \ PR\text{cd}

\quad \left[
\begin{array}{c}
(\text{"pluspp"}, P\text{Var} (\text{natT} \rightarrow \text{natT} \rightarrow \text{natT})),

(\text{"multpp"}, P\text{Var} (\text{natT} \rightarrow \text{natT} \rightarrow \text{natT}))) = R\text{cd}
\end{array}\right]

\quad \left[
\begin{array}{c}
(\text{"multpp"}, \lambda : \text{natT}. \lambda : \text{natT}. \lambda : \text{Top}. \lambda <\ : \text{TVar} \ 0. \lambda <\ : \text{TVar} \ 1. \lambda : \text{TVar} \ 2 \rightarrow \\
\text{TVar} \ 1. \ \text{Var} \ 5 \cdot \text{Var} \ 3 \cdot \text{Var} \ 2 \cdot \text{Var} \ 2 \cdot \text{Var} \ 1 \cdot \text{Var} \ 0)

(\text{"pluspp"}, \lambda : \text{natT}. \lambda : \text{natT}. \lambda : \text{Top}. \lambda <\ : \text{TVar} \ 0. \lambda <\ : \text{TVar} \ 1. \lambda : \text{TVar} \ 2 \rightarrow \\
\text{TVar} \ 1. \ \text{Var} \ 6 \cdot \\
\text{Var} \ 4 \cdot \\
\text{Var} \ 3 \cdot \\
\text{Var} \ 2 \cdot \\
\text{Var} \ 1 \cdot \\
\text{Var} \ 5 \cdot \\
\text{Var} \ 4 \cdot \\
\text{Var} \ 3 \cdot \\
\text{Var} \ 2 \cdot \\
\text{Var} \ 1 \cdot \text{Var} \ 0))
\end{array}\right]

\text{IN}

\text{Var} \ 0 \cdot (\text{Var} \ 1 \cdot \text{Var} \ 2 \cdot \text{Var} \ 2) \cdot \text{Var} \ 2

\textbf{value} normal-forms \ \text{fact2}

Unfortunately, the definition based on evaluation contexts from \S4 is not directly executable. The reason is that from the definition of evaluation contexts, the code generator cannot immediately read off an algorithm that, given a term \(t\), computes a context \(E\) and a term \(t_0\) such that \(t = E \ t_0\). In order to do this, one would have to extract the algorithm contained in the proof of the \textit{decomposition lemma} from \S4.
values \{ u. \text{norm} \text{fact2} u \}

References

