Practical type inference for arbitrary-rank types
Technical Appendix

Dimitrios Vytiniotis   Stephanie Weirich   Simon Peyton-Jones
University of Pennsylvania   Microsoft Research
{dimitriv,sweirich}@cis.upenn.edu   simonpj@microsoft.com

September 27, 2005

Note: This document accompanies the paper "Practical type inference for arbitrary-rank types" [6]. Prior reading of the main paper is required.

Contents

List of Figures

1 Introduction

This document is structured as follows: We first study several formalisations of polymorphic subsumption relations in Section 2. In Section 3 we give the most interesting properties of several type systems for higher-rank types, including the Odersky-Läufer type system [5], and study the connection between them and between the original Damas-Milner type system. We specifically focus on the bidirectional higher-rank type system, which is the main type system of the paper "Practical type inference for arbitrary-rank types". Finally, in Section 4 we give the formalisation of a sound and complete algorithm for the bidirectional type system. The algorithm is a straightforward extension of "Algorithm W" [3, 1].

The language that we use throughout the document is given in Figures 1 and 2. Our notation is standard. We use $\mathcal{S}$, $\mathcal{P}$, $\mathcal{T}$ for the sets of σ, ρ, and τ-types respectively. Substitutions, denoted with S, T, U, V are, as usual, idempotent finite maps from variables to **monotypes**. We use $dom(S)$ and $range(S)$ to denote the domain and the range of a substitution S respectively. We define $S(a) = a$ whenever $a \notin dom(S)$. Overloading the notation, we write $\overline{a} \notin \overline{b}$ to mean that the two sets of variables are **disjoint**; moreover for two sets of variables $\mathcal{X}_1$ and $\mathcal{X}_2$ we write $\mathcal{X}_1, \mathcal{X}_2$ to denote their **union**. Composition of substitutions, $S \cdot V$, is defined as usual: $S \cdot V(\sigma) = S(V(\sigma))$. A comprehensive account of substitutions and their algebraic properties is beyond the scope of this document and can be found elsewhere, for example in [2].

$$
\begin{array}{llll}
t, u & ::= & i & \text{integer literal} \\
     & \mid & x & \text{variable} \\
     & \mid & \backslash x \,.\, t & \text{abstraction} \\
     & \mid & \backslash (x :: \sigma) \,.\, t & \text{annotated abstraction} \\
     & \mid & t\ u & \text{application} \\
     & \mid & \texttt{let } x \texttt{ = } u \texttt{ in } t & \text{let generalisation} \\
     & \mid & t :: \sigma & \text{annotated term}
\end{array}
$$

Figure 1: Syntax of terms

2 Polymorphic subsumption relations

In this section we study the relations given in Figure 3, Figure 4, Figure 5, and Figure 6. We give transitivity, reflexivity and substitution lemmas for all relations and we associate each other. We discuss the properties and three different formalisations of the predicative fragment of Mitchell's F-eta containment relation [4].

2.1 Odersky-Läufer subsumption

The Odersky-Läufer subsumption relation is given in Figure 3.

Lemma 2.1 (Substitution). *If $\vdash^{ol} \sigma_1 \leq \sigma_2$ then $\vdash^{ol} S\sigma_1 \leq S\sigma_2$, and the new derivation has the same height.*

$$
\begin{array}{llll}
\sigma & ::= & \forall \overline{a} \,.\, \rho & \text{polytypes} \\
\rho & ::= & \tau \mid \sigma \to \sigma & \\
\tau & ::= & a \mid \tau \to \tau \mid \texttt{Int} & \text{monotypes}
\end{array}
$$

Figure 2: Syntax of types

$$\boxed{\vdash^{ol} \sigma \leq \sigma'}$$

$$\dfrac{\overline{a} \notin \mathit{ftv}(\sigma) \quad \vdash^{ol} \sigma \leq \rho}{\vdash^{ol} \sigma \leq \forall \overline{a}.\rho} \ \text{SKOL} \qquad\qquad \dfrac{\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2}{\vdash^{ol} \forall \overline{a}.\rho_1 \leq \rho_2} \ \text{SPEC}$$

$$\dfrac{\vdash^{ol} \sigma_3 \leq \sigma_1 \quad \vdash^{ol} \sigma_2 \leq \sigma_4}{\vdash^{ol} (\sigma_1 \rightarrow \sigma_2) \leq (\sigma_3 \rightarrow \sigma_4)} \ \text{FUN} \qquad \dfrac{}{\vdash^{ol} \tau \leq \tau} \ \text{MONO}$$

Figure 3: Subsumption in the Odersky-Läufer type system

Proof. By induction on the height of the derivation. We proceed by case analysis on the last rule used.

- Case SKOL. In this case we have that $\sigma_2 = \forall \overline{a}.\rho_2$, given that $\overline{a} \notin \mathit{ftv}(\sigma_1)$ and $\vdash^{ol} \sigma_1 \leq \rho_2$. Consider the substitution $S \cdot [\overline{a \mapsto b}]$ where $\overline{b} \notin \mathit{ftv}(\sigma_1), \overline{a}, \mathit{vars}(S)$. Then, by induction hypothesis: $\vdash^{ol} S[\overline{a \mapsto b}]\sigma_1 \leq S[\overline{a \mapsto b}]\rho_2$, or equivalently, $\vdash^{ol} S\sigma_1 \leq S[\overline{a \mapsto b}]\rho_2$. But now we can apply rule SKOL to get $\vdash^{ol} S\sigma_1 \leq \forall \overline{b}.S[\overline{a \mapsto b}]\rho_2$, or using α-renaming, $\vdash^{ol} S\sigma_1 \leq S(\forall \overline{a}.\rho_2)$.

- Case SPEC. In this case $\sigma_1 = \forall \overline{a}.\rho$, $\sigma_2 = \rho_2$, and by the premises of the rule there exist some $\overline{\tau}$ such that $\vdash^{ol} [\overline{a \mapsto \tau}]\rho \leq \rho_2$. We need to show that $\vdash^{ol} S(\forall \overline{a}.\rho) \leq S\rho_2$, or $\vdash^{ol} \forall \overline{b}.S[\overline{a \mapsto b}]\rho \leq S\rho_2$ for fresh $\overline{b} \notin \mathit{vars}(S), \mathit{ftv}(\overline{\tau})$. By rule SPEC it is enough to find types $\overline{\tau'}$ such that $\vdash^{ol} [\overline{b \mapsto \tau'}]S[\overline{a \mapsto b}]\rho \leq S\rho_2$, and $\overline{b} \notin \mathit{vars}(S)$. Pick $\overline{\tau'} = \overline{S\tau}$. Then we have to show that: $\vdash^{ol} [\overline{b \mapsto S\tau}]S[\overline{a \mapsto b}]\rho \leq S\rho_2$, or equivalently $\vdash^{ol} S[\overline{b \mapsto \tau}][\overline{a \mapsto b}]\rho \leq S\rho_2$, or equivalently $\vdash^{ol} S[\overline{a \mapsto \tau}]\rho \leq S\rho_2$ and this follows by induction hypothesis.

- Case FUN. Follows by the induction hypotheses and rule FUN.

- Case MONO. Trivial.

$\square$

Lemma 2.2 (Reflexivity). $\vdash^{ol} \sigma \leq \sigma$.

Proof. By induction on the size of σ. We proceed by case analysis on the structure of σ. The case when $\sigma = \tau$ follows by rule MONO. The case when $\sigma = \sigma_1 \rightarrow \sigma_2$ follows by rule FUN since, by induction hypothesis, $\vdash^{ol} \sigma_1 \leq \sigma_1$ and $\vdash^{ol} \sigma_2 \leq \sigma_2$. For the case when $\sigma = \forall \overline{a}.\rho$, we know by induction hypothesis that $\vdash^{ol} \rho \leq \rho$ and the result follows by an application of SPEC and SKOL. $\square$

Lemma 2.3 (Transitivity). *If* $\vdash^{ol} \sigma_1 \leq \sigma_2$ *and* $\vdash^{ol} \sigma_2 \leq \sigma_3$ *then* $\vdash^{ol} \sigma_1 \leq \sigma_3$.

Proof. We prove the lemma by induction on the sums of heights of the two derivations. We proceed by case analysis on the last rule used in each derivation. We have the following combinations for the last rule of the first and the last rule of the second derivation.

- Case SKOL-SKOL. In this case $\sigma_2 = \forall \overline{a}.\rho_2$ and $\sigma_3 = \forall \overline{b}.\rho_3$. By the premises of the first derivation we have $\vdash^{ol} \sigma_1 \leq \rho_2$ and $\overline{a} \notin \mathit{ftv}(\sigma_1)$. By the premises of the second derivation we have

$$\vdash^{ol} \forall \overline{a}.\rho_2 \leq \rho_3 \tag{1}$$

5

and $\overline{b} \notin \mathit{ftv}(\forall \overline{a} . \rho_2)$. Consider a substitution $[\overline{b \to c}]$ with $\overline{c} \notin \mathit{ftv}(\sigma_1, \sigma_2, \sigma_3)$. Then by the substitution lemma we get $\vdash^{ol} \forall \overline{a} . \rho_2 \leq [\overline{b \mapsto c}]\rho_3$. and this derivation has the same height as (1). Then we can apply the induction hypothesis to get that $\vdash^{ol} \sigma_1 \leq [\overline{b \mapsto c}]\rho_3$ and by rule SKOL we get $\vdash^{ol} \sigma_1 \leq \forall \overline{c} . [\overline{b \mapsto c}]\rho_3$. Using α-renaming, $\vdash^{ol} \sigma_1 \leq \forall \overline{b} . \rho_3$.

- Case SKOL-SPEC. Here $\sigma_2 = \forall \overline{a} . \rho_2$ and $\sigma_3 = \rho_3$. Then we have $\vdash^{ol} \sigma_1 \leq \forall \overline{a} . \rho_2$ given that

$$\vdash^{ol} \sigma_1 \leq \rho_2 \tag{2}$$
$$\overline{a} \notin \mathit{ftv}(\sigma_1) \tag{3}$$

Additionally, for some $\overline{\tau}$,

$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_2 \leq \rho_3 \tag{4}$$

By the substitution lemma $\vdash^{ol} [\overline{a \mapsto \tau}]\sigma_1 \leq [\overline{a \mapsto \tau}]\rho_2$ and the derivation has the same height as (2). By (3) it follows that $\vdash^{ol} \sigma_1 \leq [\overline{a \mapsto \tau}]\rho_2$ and by induction hypothesis, using (4), $\vdash^{ol} \sigma_1 \leq \rho_3$.

- Case SKOL-FUN/MONO. Can't happen.

- Case SPEC-SKOL. Here $\sigma_1 = \forall \overline{a} . \rho_1$, $\sigma_2 = \rho_2$ for some ρ_2, and $\sigma_3 = \forall \overline{b} . \rho_3$. By the premises of the first derivation

$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2 \tag{5}$$

for some $\overline{\tau}$. By the premises of the second derivation[1] we get

$$\vdash^{ol} \rho_2 \leq \rho_3 \tag{6}$$

and $\overline{b} \notin \mathit{ftv}(\rho_2)$. Consider a renaming substitution $[\overline{b \mapsto c}]$, such that $\overline{c} \notin \mathit{ftv}(\sigma_1, \sigma_2, \sigma_3)$. Then by the substitution lemma and (6) we get $\vdash^{ol} \rho_2 \leq [\overline{b \mapsto c}]\rho_3$ with the same height. By this, (5), and induction hypothesis $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq [\overline{b \mapsto c}]\rho_3$. By rule SPEC we get $\vdash^{ol} \forall \overline{a} . \rho_1 \leq [\overline{b \mapsto c}]\rho_3$ and by rule SKOL, $\vdash^{ol} \forall \overline{a} . \rho_1 \leq \forall \overline{c} . [\overline{b \mapsto c}]\rho_3$. With an α-renaming $\vdash^{ol} \forall \overline{a} . \rho_1 \leq \forall \overline{b} . \rho_3$.

- Case SPEC-SPEC. Can't happen.

- Case SPEC-FUN. We have that $\sigma_1 = \forall \overline{a} . \rho_1$, and that $\sigma_2 = \sigma_{21} \to \sigma_{22}$. By the premises we have that, for some $\overline{\tau}$, $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \sigma_2$. Therefore by induction hypothesis $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \sigma_3$ and by applying rule SPEC we are done.

- Case SPEC-MONO. Easy.

- Case FUN-SKOL. Here $\sigma_1 = \sigma_{11} \to \sigma_{12}$, $\sigma_2 = \sigma_{21} \to \sigma_{22}$, and $\sigma_3 = \forall \overline{a} . \rho_3$. Moreover

$$\vdash^{ol} \sigma_2 \leq \rho_3 \tag{7}$$

where $\overline{a} \notin \mathit{ftv}(\sigma_2)$. Consider a renaming substitution $[\overline{a \mapsto b}]$, such that $\overline{b} \notin \mathit{ftv}(\sigma_1, \sigma_2, \sigma_3)$. By the substitution lemma $\vdash^{ol} \sigma_2 \leq [\overline{a \mapsto b}]\rho_3$ and this derivation has the same height as (7). Then by induction hypothesis $\vdash^{ol} \sigma_1 \leq [\overline{a \mapsto b}]\rho_3$ and by rule SKOL we get $\vdash^{ol} \sigma_1 \leq \forall \overline{b} . [\overline{a \mapsto b}]\rho_3$, or with an α-renaming, $\vdash^{ol} \sigma_1 \leq \forall \overline{a} . \rho_3$.

- Case FUN-SPEC. Can't happen.

- Case FUN-FUN. In this case we have that $\sigma_1 = \sigma_{11} \to \sigma_{12}$ and $\sigma_2 = \sigma_{21} \to \sigma_{22}$ and $\sigma_3 = \sigma_{31} \to \sigma_{32}$. Moreover $\vdash^{ol} \sigma_{31} \leq \sigma_{21}$ and $\vdash^{ol} \sigma_{21} \leq \sigma_{11}$ and by induction hypothesis we get that $\vdash^{ol} \sigma_{31} \leq \sigma_{11}$. Also we have that $\vdash^{ol} \sigma_{12} \leq \sigma_{22}$ and $\vdash^{ol} \sigma_{22} \leq \sigma_{32}$. Then, by induction hypothesis $\vdash^{ol} \sigma_{12} \leq \sigma_{32}$ and by applying rule FUN we are done.

---

[1] Notice that it is not the case that $\overline{b}$ must be empty; the reason is that our types are not in prenex form. Consider for example the subsumption check $\vdash^{ol} \mathtt{Int} \to \forall a . a \to a \leq \forall c . \mathtt{Int} \to c \to c$.

$$\boxed{\overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma \leq \sigma'}$$

$$\frac{\overline{\sigma}, \sigma_2 \vdash^{\mathfrak{Dsk}} \sigma_1 \leq \sigma_3}{\overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_1 \leq (\sigma_2 \to \sigma_3)} \text{ RFUN} \qquad \frac{\overline{a} \notin \mathit{ftv}(\sigma_1, \overline{\sigma}) \quad \overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_1 \leq \rho}{\overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_1 \leq \forall \overline{a}.\rho} \text{ SKOL}$$

$$\frac{\vdash^{\mathfrak{Dsk}} \sigma_3 \leq \sigma_1 \quad \overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_2 \leq \tau}{\sigma_3, \overline{\sigma} \vdash^{\mathfrak{Dsk}} (\sigma_1 \to \sigma_2) \leq \tau} \text{ LFUN} \qquad \frac{\overline{\sigma} \vdash^{\mathfrak{Dsk}} [\overline{a \mapsto \tau}] \rho \leq \tau_1}{\overline{\sigma} \vdash^{\mathfrak{Dsk}} \forall \overline{a}.\rho \leq \tau_1} \text{ SPEC}$$

$$\frac{}{\vdash^{\mathfrak{Dsk}} \tau \leq \tau} \text{ MONO}$$

Figure 4: Subsumption with deep skolemisation

- Case FUN-MONO. Easy.

- Case MONO-SPEC. Can't happen.

- Case MONO-SKOL/FUN/MONO. Trivial.

$\square$

The Odersky-Läufer subsumption is syntax-directed, and therefore has nice inversion properties. The following lemmas capture inversion.

Lemma 2.4 (Skolemisation inversion). *If $\overline{a} \notin \mathit{ftv}(\sigma)$ and $\vdash^{ol} \sigma \leq \forall \overline{a}.\rho$, then $\vdash^{ol} \sigma \leq \rho$.*

Proof. Straightforward induction. $\square$

Lemma 2.5 (Specialisation inversion). *If $\vdash^{ol} \forall \overline{a}.\rho_1 \leq \rho_2$, then $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2$ for some $\overline{\tau}$.*

Proof. Straightforward induction. $\square$

Lemma 2.6 (Fun inversion). *If $\vdash^{ol} \rho_1 \leq \sigma_3 \to \sigma_4$, then $\rho_1 = \sigma_1 \to \sigma_2$ with $\vdash^{ol} \sigma_3 \leq \sigma_1$ and $\vdash^{ol} \sigma_2 \leq \sigma_4$.*

Proof. Straightforward induction. $\square$

2.2 Deep skolemisation subsumption—sequent-style

In Figure 4 we give a relation that performs the "skolemisation" step deeply to the right of arrow types. It resembles a sequent-style presentation. Here are the most important properties about this relation.

Lemma 2.7 (Substitution). *If $\overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_1 \leq \sigma_2$ then $S\overline{\sigma} \vdash^{\mathfrak{Dsk}} S\sigma_1 \leq S\sigma_2$, and the new derivation has the same height.*

Proof. By induction on the height of the derivation. We proceed by case analysis on the last rule used.

- Case SKOL. In this case we have that $\sigma_2 = \forall \overline{a}.\rho$, given that $\overline{a} \notin \mathit{ftv}(\sigma_1, \overline{\sigma})$ and $\overline{\sigma} \vdash^{\mathfrak{Dsk}} \sigma_1 \leq \rho$. Consider the substitution $S \cdot [\overline{a \mapsto b}]$ where $\overline{b} \notin \mathit{ftv}(\sigma_1, \overline{\sigma})$ and $\overline{b} \notin \overline{a}, \mathit{vars}(S)$. Then, by induction hypothesis: $S[\overline{a \mapsto b}]\overline{\sigma} \vdash^{\mathfrak{Dsk}} S[\overline{a \mapsto b}]\sigma_1 \leq S[\overline{a \mapsto b}]\rho$, or equivalently, $S\overline{\sigma} \vdash^{\mathfrak{Dsk}} S\sigma_1 \leq S[\overline{a \mapsto b}]\rho$. But now we can apply rule SKOL to get $S\overline{\sigma} \vdash^{\mathfrak{Dsk}} S\sigma_1 \leq \forall \overline{b}.S[\overline{a \mapsto b}]\rho$, or, using α-renaming $S\overline{\sigma} \vdash^{\mathfrak{Dsk}} S\sigma_1 \leq S(\forall \overline{a}.\rho)$.

7

- Case SPEC. In this case $\sigma_1 = \forall\overline{a}.\rho$, $\sigma_2 = \tau_2$, and by the premises of the rule there exist some $\overline{\tau}$ such that $\overline{\sigma} \vdash^{\partial s\ell} [\overline{a \mapsto \tau}]\rho \leq \tau_2$. We need to show that $S\overline{\sigma} \vdash^{\partial s\ell} S(\forall\overline{a}.\rho) \leq S\tau_2$, or with an α-renaming, $S\overline{\sigma} \vdash^{\partial s\ell} \forall\overline{b}.S[\overline{a \mapsto b}]\rho \leq S\tau_2$ for $\overline{b} \notin vars(S), ftv(\overline{\tau}, \overline{\sigma})$. By rule SPEC it is enough to find types $\overline{\tau'}$ such that $S\overline{\sigma} \vdash^{\partial s\ell} [\overline{b \mapsto \tau'}]S[\overline{a \mapsto b}]\rho \leq S\tau_2$. Pick $\overline{\tau'} = \overline{S\tau}$. Then it remains to show that $S\overline{\sigma} \vdash^{\partial s\ell} [\overline{b \mapsto S\tau}]S[\overline{a \mapsto b}]\rho \leq S\tau_2$, or $S\overline{\sigma} \vdash^{\partial s\ell} S[\overline{b \mapsto \tau}][\overline{a \mapsto b}]\rho \leq S\tau_2$, or $S\overline{\sigma} \vdash^{\partial s\ell} S[\overline{a \mapsto \tau}]\rho \leq S\tau_2$, but this holds by induction hypothesis.

- Case RFUN. Follows by induction hypothesis and application of rule RFUN.

- Case LFUN. Follows by induction hypothesis and application of rule LFUN.

- Case MONO. Trivial.

$\square$

For the rest of this section we are going to give the connection between the Odersky-Läufer subsumption and the deep skolemisation subsumption. Namely we show that two types are related in the deep skolemisation subsumption iff their prenex forms are related in Odersky-Läufer.

Definition 2.8 (Prenex conversion). *The function* $\mathfrak{pr}(\cdot) : \mathcal{S} \to \mathcal{S}$ *is defined as follows:*

$$
\mathfrak{pr}(\sigma) = \begin{cases}
\tau & \sigma = \tau \\
\forall\overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2 & \sigma = \sigma_1 \to \sigma_2 \wedge \sigma \neq \tau \\
& \mathfrak{pr}(\sigma_2) = \forall\overline{a}.\rho_2 \\
& \overline{a} \notin ftv(\sigma_1) \\
\forall\overline{a}\overline{b}.\rho_2 & \sigma = \forall\overline{a}.\rho_1 \wedge \overline{a} \neq \emptyset \\
& \mathfrak{pr}(\rho_1) = \forall\overline{b}.\rho_2 \\
& \overline{b} \notin \overline{a}
\end{cases}
$$

Definition 2.9 (Function conversion). *The function* $fun(\cdot\,;\cdot) : (\overline{\mathcal{S}} \times \mathcal{S}) \to \mathcal{S}$ *is defined as follows:*

$$
fun(\overline{\sigma}\,;\sigma) \begin{cases}
\sigma & \overline{\sigma} = \epsilon \\
fun(\overline{\sigma}_1\,;\sigma_1 \to \sigma) & \overline{\sigma} = \overline{\sigma}_1, \sigma_1
\end{cases}
$$

We naturally extend the $\mathfrak{pr}(\cdot)$ function for sequences of types, by mapping the prenex function across every type in the sequence. The next lemmas give the algebraic properties of th $\mathfrak{pr}(\cdot)$ and $fun(\cdot\,;\cdot)$ functions.

Lemma 2.10 (Prenex conversion preserves size). $size(\sigma) = size(\mathfrak{pr}(\sigma))$.

Proof. By structural induction on the type σ. If $\sigma = \tau$ then $size(\mathfrak{pr}(\tau)) = size(\tau)$ by definition. If $\sigma = \sigma_1 \to \sigma_2$, not a monotype, then $\mathfrak{pr}(\sigma_1 \to \sigma_2) = \forall\overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2$, where $\forall\overline{a}.\rho_2 = \mathfrak{pr}(\sigma_2)$, thus $size(\mathfrak{pr}(\sigma_2)) =\mid \overline{a} \mid +size(\rho_2)$. By induction $size(\sigma_2) =\mid \overline{a} \mid +size(\rho_2)$. By induction $size(\sigma_1) = size(\mathfrak{pr}(\sigma_1))$. Then

$$
\begin{aligned}
size(\sigma_1 \to \sigma_2) &= size(\sigma_1) + size(\sigma_2) + 1 \\
&= size(\mathfrak{pr}(\sigma_1)) + \mid \overline{a} \mid +size(\rho_2) + 1 \\
&= size(\forall\overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2)
\end{aligned}
$$

$\square$

Lemma 2.11. $fun(\sigma_1, \overline{\sigma}\,;\sigma) = \sigma_1 \to fun(\overline{\sigma}\,;\sigma)$.

Proof. By induction on the size of $\overline{\sigma}$. If $\overline{\sigma} = \epsilon$ then

$$
\begin{aligned}
fun(\sigma_1\,;\sigma) &= fun(\epsilon\,;\sigma_1 \to \sigma) \\
&= \sigma_1 \to \sigma \\
&= \sigma_1 \to fun(\epsilon\,;\sigma)
\end{aligned}
$$

8

If $\overline{\sigma} = \overline{\sigma}_2, \sigma_2$ then

$$
\begin{aligned}
fun(\sigma_1, \overline{\sigma}_2, \sigma_2 \,;\, \sigma) &= fun(\sigma_1, \overline{\sigma}_2 \,;\, \sigma_2 \to \sigma) \\
&= \sigma_1 \to fun(\overline{\sigma}_2 \,;\, \sigma_2 \to \sigma) \quad \text{(by induction)} \\
&= \sigma_1 \to fun(\overline{\sigma}_2, \sigma_2 \,;\, \sigma) \\
&= \sigma_1 \to fun(\overline{\sigma} \,;\, \sigma)
\end{aligned}
$$

$\square$

Lemma 2.12. $\mathfrak{pr}(\mathfrak{pr}(\sigma)) = \mathfrak{pr}(\sigma)$.

Proof. By induction on the size of σ. If $\sigma = \tau$ then it is trivial. If $\sigma = \sigma_1 \to \sigma_2$, where $\sigma \neq \tau$, then $\mathfrak{pr}(\mathfrak{pr}(\sigma_1 \to \sigma_2)) = \mathfrak{pr}(\forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2)$, where $\mathfrak{pr}(\sigma_2) = \forall \overline{a}.\rho_2$ and $\overline{a} \notin ftv(\sigma_1)$. By induction hypothesis, we get:

$$\mathfrak{pr}(\forall \overline{a}.\rho_2) = \forall \overline{a}.\rho_2 \tag{1}$$

$$\mathfrak{pr}(\mathfrak{pr}(\sigma_1)) = \mathfrak{pr}(\sigma_1) \tag{2}$$

Now we have two cases:

- $\overline{a} = \emptyset$. In this case $\mathfrak{pr}(\forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2) = \mathfrak{pr}(\mathfrak{pr}(\sigma_1) \to \rho_2)$. By (1), $\mathfrak{pr}(\rho_2) = \rho_2$ and using (2) and the definition $\mathfrak{pr}(\mathfrak{pr}(\sigma_1) \to \rho_2) = \mathfrak{pr}(\sigma_1) \to \rho_2$.

- $\overline{a} \neq \emptyset$. In order to compute $\mathfrak{pr}(\forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2)$ we have to compute $\mathfrak{pr}(\mathfrak{pr}(\sigma_1) \to \rho_2)$. But $size(\mathfrak{pr}(\sigma_1) \to \rho_2) = size(\mathfrak{pr}(\sigma_1)) + size(\rho_2) + 1$ and using Lemma 2.10 we have that $size(\mathfrak{pr}(\sigma_1) \to \rho_2) = size(\sigma_1) + size(\rho_2) + 1$ and since $\overline{a} \neq \emptyset$ this is less than $size(\sigma)$. Then by induction hypothesis we get $\mathfrak{pr}(\mathfrak{pr}(\sigma_1) \to \rho_2) = \mathfrak{pr}(\sigma_1) \to \rho_2$. Using the definition $\mathfrak{pr}(\forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2) = \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2$.

Finally if $\sigma = \forall \overline{a}.\rho$ for $\overline{a} \neq \emptyset$ we have that $\mathfrak{pr}(\forall \overline{a}.\rho) = \forall \overline{a}\overline{b}.\rho_1$, where $\forall \overline{b}.\rho_1 = \mathfrak{pr}(\rho)$. Now, by induction, it must be that $\mathfrak{pr}(\forall \overline{b}.\rho_1) = \forall \overline{b}.\rho_1$. If $\overline{b} = \emptyset$ then $\mathfrak{pr}(\rho_1) = \rho_1$. If $\overline{b} \neq \emptyset$ then by definition of prenex conversion if must be that $\mathfrak{pr}(\forall \overline{b}.\rho_1) = \forall \overline{b}\overline{c}.\rho_2$ where $\mathfrak{pr}(\rho_1) = \forall \overline{c}.\rho_2$. But this means that $\overline{c} = \emptyset$ and $\rho_2 = \rho_1$, therefore in every case $\mathfrak{pr}(\rho_1) = \rho_1$. Using this and the definition of prenex conversion $\mathfrak{pr}(\mathfrak{pr}(\forall \overline{a}.\rho)) = \mathfrak{pr}(\forall \overline{a}\overline{b}.\rho_1) = \forall \overline{a}\overline{b}.\rho_1$. $\square$

Corollary 2.13. *If* $\mathfrak{pr}(\sigma) = \forall \overline{a}.\rho$ *then* $\mathfrak{pr}(\rho) = \rho$.

Proof. Easy corollary of Lemma 2.12. $\square$

Lemma 2.14. *If* $\mathfrak{pr}(\sigma) = \forall \overline{a}.\rho$ *and* $\overline{a} \notin ftv(\overline{\sigma})$ *then* $\mathfrak{pr}(fun(\overline{\sigma} \,;\, \sigma)) = \forall \overline{a}.fun(\mathfrak{pr}(\overline{\sigma}) \,;\, \rho)$.

Proof. By induction on $\overline{\sigma}$. For $\overline{\sigma} = \epsilon$ we have that

$$
\begin{aligned}
\mathfrak{pr}(fun(\epsilon \,;\, \sigma)) &= \mathfrak{pr}(\sigma) \\
&= \forall \overline{a}.\rho \\
&= \forall \overline{a}.fun(\mathfrak{pr}(\epsilon) \,;\, \rho)
\end{aligned}
$$

For $\overline{\sigma} = \sigma_1, \overline{\sigma}_1$ we have that

$$
\begin{aligned}
\mathfrak{pr}(fun(\sigma_1, \overline{\sigma}_1 \,;\, \sigma)) &= \mathfrak{pr}(\sigma_1 \to fun(\overline{\sigma}_1 \,;\, \sigma)) \quad \text{(by Lemma 2.11)} \\
&= \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2
\end{aligned}
$$

where $\forall \overline{a}.\rho_2 = \mathfrak{pr}(fun(\overline{\sigma}_1 \,;\, \sigma))$. Hence by induction hypothesis $\rho_2 = fun(\mathfrak{pr}(\overline{\sigma}_1) \,;\, \rho)$. Therefore we get

$$
\begin{aligned}
\mathfrak{pr}(fun(\sigma_1, \overline{\sigma}_1 \,;\, \sigma)) &= \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to fun(\mathfrak{pr}(\overline{\sigma}_1 \,;\, \rho)) \\
&= \forall \overline{a}.fun(\mathfrak{pr}(\sigma_1), \mathfrak{pr}(\overline{\sigma}_1) \,;\, \rho) \quad \text{(by Lemma 2.11)} \\
&= \forall \overline{a}.fun(\mathfrak{pr}(\sigma_1, \overline{\sigma}_1) \,;\, \rho) \\
&= \forall \overline{a}.fun(\mathfrak{pr}(\overline{\sigma}) \,;\, \rho)
\end{aligned}
$$

$\square$

Lemma 2.15. *If* $\mathfrak{pr}(\sigma) = \forall \overline{a}.\rho$ *and* $\overline{a} \notin ftv(\overline{\sigma})$ *then* $\mathfrak{pr}(fun(\overline{\sigma};\sigma)) = \forall \overline{a}.\mathfrak{pr}(fun(\overline{\sigma};\rho))$ *and* $\mathfrak{pr}(fun(\overline{\sigma};\rho)) \in \mathcal{P}$.

Proof. By Lemma 2.14 it is enough to show that

$$\mathfrak{pr}(fun(\overline{\sigma};\rho)) = fun(\mathfrak{pr}(\overline{\sigma});\rho)$$

We prove this by induction on $\overline{\sigma}$. For $\overline{\sigma} = \epsilon$ we have that $\mathfrak{pr}(fun(\epsilon;\rho)) = \mathfrak{pr}(\rho)$. Using Lemma 2.13, $\mathfrak{pr}(\rho) = \rho = fun(\mathfrak{pr}(\epsilon);\rho)$. For $\overline{\sigma} = \sigma_1\overline{\sigma}_1$ we have $\mathfrak{pr}(fun(\sigma_1,\overline{\sigma}_1;\rho)) = \mathfrak{pr}(\sigma_1 \to fun(\overline{\sigma}_1;\rho))$. By induction hypothesis $\mathfrak{pr}(fun(\overline{\sigma}_1;\rho)) = fun(\mathfrak{pr}(\overline{\sigma}_1);\rho)$ Therefore

$$
\begin{aligned}
\mathfrak{pr}(\sigma_1 \to fun(\overline{\sigma}_1;\rho)) &= \mathfrak{pr}(\sigma_1) \to fun(\mathfrak{pr}(\overline{\sigma}_1);\rho) \\
&= fun(\mathfrak{pr}(\sigma_1,\overline{\sigma}_1;\rho) \\
&= fun(\mathfrak{pr}(\overline{\sigma};\rho))
\end{aligned}
$$

$\square$

Lemma 2.16 (Arrow). *If* $\vdash^{\partial s \ell} \sigma_3 \le \sigma_1$ *and* $\overline{\sigma} \vdash^{\partial s \ell} \sigma_2 \le \sigma$ *then* $\sigma_3, \overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \to \sigma_2 \le \sigma$.

Proof. By induction on the size of σ.

- Case $\sigma = \tau$. Then the assumptions match exactly the premises of rule LFUN and we are done by applying that rule.

- Case $\sigma = \sigma_a \to \sigma_b$, $\sigma \ne \tau$. Then we have that $\overline{\sigma} \vdash^{\partial s \ell} \sigma_2 \le \sigma_a \to \sigma_b$ and by easy inversion we see that $\overline{\sigma}, \sigma_a \vdash^{\partial s \ell} \sigma_2 \le \sigma_b$. Now we can apply the induction hypothesis for σ_b to get that $\sigma_3, \overline{\sigma}, \sigma_a \vdash^{\partial s \ell} \sigma_1 \to \sigma_2 \le \sigma_b$ and by applying rule RFUN we get back $\sigma_3, \sigma \vdash^{\partial s \ell} \sigma_1 \to \sigma_2 \le \sigma_a \to \sigma_b$.

- Case $\sigma = \forall \overline{a}.\rho$, $\overline{a} \ne \emptyset$. Here we have that $\overline{\sigma} \vdash^{\partial s \ell} \sigma_2 \le \forall \overline{a}.\rho$ and by inversion this can happen only using the rule SKOL. Therefore it must be that $\overline{a} \notin ftv(\sigma_2, \overline{\sigma})$ and then

$$\overline{\sigma} \vdash^{\partial s \ell} \sigma_2 \le \rho \tag{1}$$

  Consider a substitution $[\overline{a} \mapsto \overline{c}]$, such that $\overline{c} \notin ftv(\sigma_1, \sigma_3)$, $\overline{c} \notin ftv(\overline{\sigma}, \sigma_2)$. By the substitution lemma we get $\overline{\sigma} \vdash^{\partial s \ell} \sigma_2 \le [\overline{a} \mapsto \overline{c}]\rho$. Moreover $size([\overline{a} \mapsto \overline{c}]\rho) = size(\rho)$. Then, by induction hypothesis we get $\sigma_3\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \to \sigma_2 \le [\overline{a} \mapsto \overline{c}]\rho$, and by SKOL we have $\sigma_3\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \to \sigma_2 \le \forall \overline{c}.[\overline{a} \mapsto \overline{c}]\rho$. We are done with an α-renaming.

$\square$

Lemma 2.17 (Prenex recursive calls). *The number of recursive calls to* $\mathfrak{pr}(\cdot)$ *is preserved by substitution. If* $\mathfrak{pr}(\sigma)$ *uses* n *recursive calls, then so does* $\mathfrak{pr}(S\sigma)$.

Proof. By induction on σ. If $\sigma = \tau$, since $range(S) \in \mathcal{T}$ the result is trivial. Suppose $\sigma = \sigma_1 \to \sigma_2$. Then $calls(\mathfrak{pr}(\sigma_1)) = calls(\mathfrak{pr}(S\sigma_1))$ by induction hypothesis. Moreover $calls(\mathfrak{pr}(\sigma_2)) = calls(\mathfrak{pr}(S\sigma_2))$ by induction hypothesis as well and we are done. If $\sigma = \forall \overline{a}.\rho$ then, assuming without loss of generality that $\overline{a} \notin vars(S)$ we have that $calls(\mathfrak{pr}(\rho)) = calls(\mathfrak{pr}(S\rho))$ by induction hypothesis and we are done. $\square$

The next theorem says that if the prenex-canonical forms of two types are related in Odersky-Läufer subsumption, then the two types are related in $\vdash^{\partial s \ell}$.

Theorem 2.18. *If* $\mathfrak{pr}(\sigma_1) = \sigma_1'$, $\mathfrak{pr}(fun(\overline{\sigma};\sigma_2)) = \sigma_2'$, *and* $\vdash^{ol} \sigma_1' \le \sigma_2'$ *then* $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \le \sigma_2$.

Proof. By induction on the derivation $\vdash^{ol} \sigma_1' \le \sigma_2'$. We proceed by case analysis on the last rule used in the derivation.

- Case MONO. In this case we have that $\vdash^{ol} \tau \leq \tau$. By inversion it must be that $\sigma_1 = \tau$ and $fun(\overline{\sigma};\sigma_2) = \tau$, which implies that $\sigma_2 = \tau_2$ and $\overline{\sigma} = \overline{\tau}$, for some $\overline{\tau}, \tau_2$. We want to show that

$$\overline{\tau} \vdash^{\partial s \ell} fun(\overline{\tau} \,;\, \tau_2) \leq \tau_2 \tag{1}$$

We prove (1) by induction on $\overline{\tau}$.

  – Case $\overline{\tau} = \epsilon$. Then $fun(\overline{\tau} \,;\, \tau_2) = \tau_2$ and the claim follows by MONO.
  – Case $\overline{\tau} = \tau', \overline{\tau}'$. In this case by the (inner) induction hypothesis we get

$$\overline{\tau}' \vdash^{\partial s \ell} fun(\overline{\tau}' \,;\, \tau_2) \leq \tau_2 \tag{2}$$

  Moreover $\vdash^{\partial s \ell} \tau' \leq \tau'$; using this and (2) with rule LFUN we get that $\tau', \overline{\tau}' \vdash^{\partial s \ell} \tau' \to fun(\overline{\tau}';\tau_2) \leq \tau_2$, and we get (1) by this and Lemma 2.11.

- Case SKOL. Here $\mathfrak{pr}(\sigma_1) = \sigma$ and $\mathfrak{pr}(fun(\overline{\sigma} \,;\, \sigma_2)) = \forall \overline{a}.\rho$. By Lemma 2.15 it must be that $\rho = \mathfrak{pr}(fun(\overline{\sigma} \,;\, \rho_2))$ where we also assumed that $\mathfrak{pr}(\sigma_2) = \forall \overline{a}.\rho_2$ and $\overline{a} \notin ftv(\overline{\sigma})$. Moreover $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \forall \overline{a}.\mathfrak{pr}(fun(\overline{\sigma} \,;\, \rho_2))$, given that $\overline{a} \notin ftv(\mathfrak{pr}(\sigma_1)) = ftv(\sigma)$ and

$$\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(fun(\overline{\sigma} \,;\, \rho_2)) \tag{3}$$

By (3) and induction hypothesis it must be that $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \leq \rho_2$ and by applying rule SKOL we get $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \leq \forall \overline{a}.\rho_2$, since $\overline{a} \notin ftv(\overline{\sigma}, \sigma_1)$.

- Case FUN. Here we have the following

$$\mathfrak{pr}(\sigma_1) = \sigma_1' \to \sigma_2' \tag{4}$$
$$\mathfrak{pr}(fun(\overline{\sigma} \,;\, \sigma_2)) = \sigma_3' \to \sigma_4' \tag{5}$$
$$\vdash^{ol} \sigma_1' \to \sigma_2' \leq \sigma_3' \to \sigma_4' \tag{6}$$
$$\vdash^{ol} \sigma_3' \leq \sigma_1' \tag{7}$$
$$\vdash^{ol} \sigma_2' \leq \sigma_4' \tag{8}$$

By (4) it must be that

$$\sigma_1 = \sigma_{01} \to \sigma_{02} \tag{9}$$

For (5) we have two cases:

  – $\overline{\sigma} = \epsilon$. In this case it must be that $\sigma_2 = \sigma_{21} \to \sigma_{22}$. Moreover $\mathfrak{pr}(\sigma_{21}) = \sigma_3'$, $\mathfrak{pr}(\sigma_{22}) = \sigma_4'$, $\mathfrak{pr}(\sigma_{01}) = \sigma_1'$, and $\mathfrak{pr}(\sigma_{02}) = \sigma_2'$. Equivalently $\mathfrak{pr}(fun(\epsilon \,;\, \sigma_{01})) = \sigma_1'$, and by induction hypothesis

$$\vdash^{\partial s \ell} \sigma_{21} \leq \sigma_{01} \tag{10}$$

  Similarly $\mathfrak{pr}(fun(\epsilon \,;\, \sigma_{22})) = \sigma_4'$, therefore by induction hypothesis

$$\vdash^{\partial s \ell} \sigma_{02} \leq \sigma_{22} \tag{11}$$

  With (10) and (11) we can apply Lemma 2.16 to get $\sigma_{21} \vdash^{\partial s \ell} \sigma_{01} \to \sigma_{02} \leq \sigma_{22}$ and by rule RFUN $\vdash^{\partial s \ell} \sigma_{01} \to \sigma_{02} \leq \sigma_{21} \to \sigma_{22}$ as required.
  – $\overline{\sigma} = \sigma_a, \overline{\sigma}'$. Then we have that

$$\begin{aligned} \mathfrak{pr}(fun(\sigma_a, \overline{\sigma}' \,;\, \sigma_2)) &= \mathfrak{pr}(\sigma_a \to fun(\overline{\sigma}' \,;\, \sigma_2)) \\ &= \sigma_3' \to \sigma_4' \end{aligned}$$

11

Then, by definition of $\mathfrak{pr}(\cdot)$ we have

$$\mathfrak{pr}(\sigma_a) = \sigma_3' \tag{12}$$

$$\mathfrak{pr}(fun(\overline{\sigma}'\,;\sigma_2)) = \sigma_4' \tag{13}$$

By (4) and (9) we have that $\mathfrak{pr}(\sigma_{01}) = \sigma_1'$, or $\mathfrak{pr}(fun(\epsilon\,;\sigma_{01})) = \sigma_1'$. From this and (12) and induction hypothesis we get that

$$\vdash^{\partial\mathfrak{s}\ell} \sigma_a \leq \sigma_{01} \tag{14}$$

From (4) and (9) we get $\mathfrak{pr}(\sigma_{02}) = \sigma_2'$. From this and (13) and induction hypothesis we have that $\overline{\sigma}' \vdash^{\partial\mathfrak{s}\ell} \sigma_{02} \leq \sigma_2$. From this and (14) and Lemma 2.16 we get $\sigma_a\overline{\sigma}' \vdash^{\partial\mathfrak{s}\ell} \sigma_{01} \to \sigma_{02} \leq \sigma_2$, and using (9) we get $\overline{\sigma} \vdash^{\partial\mathfrak{s}\ell} \sigma_1 \leq \sigma_2$ as required.

- Case SPEC. For this case we have

$$\mathfrak{pr}(\sigma_1) = \forall\overline{a}.\rho_1 \tag{15}$$

$$\mathfrak{pr}(fun(\overline{\sigma}\,;\sigma_2)) = \rho_2 \tag{16}$$

$$\vdash^{ol} \forall\overline{a}.\rho_1 \leq \rho_2 \tag{17}$$

$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2 \tag{18}$$

We wish to show that $\overline{\sigma} \vdash^{\partial\mathfrak{s}\ell} \sigma_1 \leq \sigma_2$. We prove this claim with an **inner** induction on the number of recursive calls to $\mathfrak{pr}(\cdot)$ from $\mathfrak{pr}(\sigma_1)$. Specifically our induction hypothesis is the following:

IH: If $\mathfrak{pr}(\sigma_1^i)$ has fewer recursive calls than $\mathfrak{pr}(\sigma_1)$ then

$$\begin{array}{ll} \text{(A)} & \mathfrak{pr}(fun(\overline{\sigma}^i\,;\sigma_2^i)) = \rho_2^i \\ \text{(B)} & \mathfrak{pr}(\sigma_1^i) = \forall\overline{b}.\rho_1^i \\ \text{(C)} & \vdash^{ol} [\overline{b \mapsto \tau^i}]\rho_1^i \leq \rho_2^i \end{array} \;\Rightarrow\; \overline{\sigma}^i \vdash^{\partial\mathfrak{s}\ell} \sigma_1^i \leq \sigma_2^i$$

We proceed with a case analysis on the structure of σ_1, and without loss of generality let us assume that σ_1 is not a monotype as we would not be in the SPEC case[2].

 – Case $\sigma_1 = \sigma_{11} \to \sigma_{12}$. Then $\mathfrak{pr}(\sigma_1) = \forall\overline{a}.\mathfrak{pr}(\sigma_{11}) \to \rho_{12}$ where

$$\mathfrak{pr}(\sigma_{12}) = \forall\overline{a}.\rho_{12} \tag{19}$$

and $\overline{a} \notin ftv(\sigma_{11})$. By (18) it must be that $\vdash^{ol} [\overline{a \mapsto \tau}](\mathfrak{pr}(\sigma_{11}) \to \rho_{12}) \leq \rho_2$, or $\vdash^{ol} \mathfrak{pr}(\sigma_{11}) \to [\overline{a \mapsto \tau}]\rho_{12} \leq \rho_2$. By inversion it must be that $\rho_2 = \sigma_2^A \to \sigma_2^B$ and

$$\vdash^{ol} \sigma_2^A \leq \mathfrak{pr}(\sigma_{11}) \tag{20}$$

$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_{12} \leq \sigma_2^B \tag{21}$$

By (16) then it must be that

$$\mathfrak{pr}(fun(\overline{\sigma}\,;\sigma_2)) = \sigma_2^A \to \sigma_2^B \tag{22}$$

We continue with case analysis on $\overline{\sigma}$.

---

[2]To be precise, if the case was a trivial application of SPEC we could just appeal to the (outer) induction hypothesis to get the result.

* Case $\overline{\sigma} = \epsilon$. Then it must be the case that $\sigma_2 = \sigma_{21} \to \sigma_{22}$, such that $\mathfrak{pr}(\sigma_{21}) = \sigma_2^A$ and $\mathfrak{pr}(\sigma_{22}) = \sigma_2^B$. By (20) we have that $\mathfrak{pr}(\sigma_{21}) \le \mathfrak{pr}(\sigma_{11})$ and by the (outer) induction hypothesis

$$\vdash^{\partial s \ell} \sigma_{21} \le \sigma_{11} \tag{23}$$

  Similarly by (21) we have that $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_{12} \le \mathfrak{pr}(\sigma_{22})$ and by an application of SPEC we have $\vdash^{ol} \forall \overline{a} . \rho_{12} \le \mathfrak{pr}(\sigma_{22})$. But the height of this derivation is still one less than the height of the derivation we are examining and therefore by (outer) induction hypothesis we have that

$$\vdash^{\partial s \ell} \sigma_{12} \le \sigma_{22} \tag{24}$$

  From (23) and (24) and Lemma 2.16 we get that $\sigma_{21} \vdash^{\partial s \ell} \sigma_{11} \to \sigma_{12} \le \sigma_{22}$ and by applying RFUN we are done.

* Case $\overline{\sigma} = \sigma_0, \overline{\sigma}_0$. In this case $\mathfrak{pr}(fun(\overline{\sigma} ; \sigma_2)) = \mathfrak{pr}(\sigma_0 \to fun(\overline{\sigma}_0 ; \sigma_2))$ which implies that

$$\mathfrak{pr}(\sigma_0) = \sigma_2^A \tag{25}$$
$$\mathfrak{pr}(fun(\overline{\sigma}_0 ; \sigma_2)) = \sigma_2^B \tag{26}$$

  From (26), (19), and (21) and the (inner) induction hypothesis (for σ_{12}) we get that

$$\overline{\sigma}_0 \vdash^{\partial s \ell} \sigma_{12} \le \sigma_2 \tag{27}$$

  From (25) and (20) and the outer induction hypothesis we get

$$\vdash^{\partial s \ell} \sigma_0 \le \sigma_{11} \tag{28}$$

  From (27) and (28) and Lemma 2.16 we get $\sigma_0, \overline{\sigma}_0 \vdash^{\partial s \ell} \sigma_{11} \to \sigma_{12} \le \sigma_2$, or equivalently $\overline{\sigma} \vdash^{\partial s \ell} \sigma_{11} \to \sigma_{12} \le \sigma_2$.

– Case $\sigma_1 = \forall \overline{a}_1 . \rho_{11}$. Here we have that $\mathfrak{pr}(\sigma_1) = \forall \overline{a}_1 \overline{a}_2 . \rho_{12}$, such that $\forall \overline{a}_2 . \rho_{12} = \mathfrak{pr}(\rho_{11})$. From this it follows that

$$\forall \overline{a}_2 . [\overline{a_1 \to \tau_1}]\rho_{12} = \mathfrak{pr}([\overline{a_1 \to \tau_1}]\rho_{11}) \tag{29}$$

  From Lemma 2.17 we know that the number of recursive calls of $\mathfrak{pr}([\overline{a_1 \to \tau_1}]\rho_{11})$ is the same as that of $\mathfrak{pr}(\rho_{11})$. We also know that

$$\vdash^{ol} [\overline{a_2 \to \tau_2}][\overline{a_1 \mapsto \tau_1}]\rho_{12} \le \rho_2 \tag{30}$$

  Then we can apply the inner induction hypothesis to get that $\overline{\sigma} \vdash^{\partial s \ell} [\overline{a_1 \to \tau_1}]\rho_{11} \le \rho_2$ and by applying rule SPEC $\overline{\sigma} \vdash^{\partial s \ell} \forall \overline{a}_1 . \rho_{11} \le \rho_2$, as required.

$\square$

Theorem 2.19. *If $\mathfrak{pr}(\sigma_1) = \sigma_1'$, $\mathfrak{pr}(fun(\overline{\sigma} ; \sigma_2)) = \sigma_2'$, and $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \le \sigma_2$ then $\vdash^{ol} \sigma_1' \le \sigma_2'$.*

Proof. By induction on the derivation $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \le \sigma_2$. We proceed by case analysis on the last rule used.

* Case MONO. In this case $\overline{\sigma} = \epsilon$, $\sigma_1 = \tau = \sigma_2$, $\mathfrak{pr}(\sigma_1) = \tau$, $\mathfrak{pr}(fun(\epsilon ; \tau)) = \tau$ and the result follows by rule MONO.

* Case RFUN. Here $\overline{\sigma} \vdash^{\partial s \ell} \sigma_1 \le \sigma_2 \to \sigma_3$ given that $\overline{\sigma}, \sigma_2 \vdash^{\partial s \ell} \sigma_1 \le \sigma_3$. Then $\mathfrak{pr}(fun(\overline{\sigma} ; \sigma_2 \to \sigma_3)) = \mathfrak{pr}(fun(\overline{\sigma}, \sigma_2 ; \sigma_3))$ and by induction hypothesis we know that $\vdash^{ol} \sigma_1' \le \mathfrak{pr}(fun(\overline{\sigma}, \sigma_2 ; \sigma_3))$.

- Case LFUN. In this case we have that

$$\sigma_3, \overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \to \sigma_2 \le \tau \tag{1}$$

$$\vdash^{\partial s\ell} \sigma_3 \le \sigma_1 \tag{2}$$

$$\overline{\sigma} \vdash^{\partial s\ell} \sigma_2 \le \tau \tag{3}$$

Moreover we know that $\mathfrak{pr}(\sigma_1 \to \sigma_2) = \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2$, with $\forall \overline{a}.\rho_2 = \mathfrak{pr}(\sigma_2)$ and $\overline{a} \notin ftv(\sigma_1)$. On the other hand $\mathfrak{pr}(fun(\sigma_3, \overline{\sigma}\,;\tau)) = \mathfrak{pr}(\sigma_3) \to \mathfrak{pr}(fun(\overline{\sigma}\,;\tau))$ by Lemma 2.15. So we need to show that $\vdash^{ol} \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2 \le \mathfrak{pr}(\sigma_3) \to \mathfrak{pr}(fun(\overline{\sigma}\,;\tau))$. By rule SPEC it is enough to show that $\vdash^{ol} \mathfrak{pr}(\sigma_1) \to [\overline{a \mapsto \tau}]\rho_2 \le \mathfrak{pr}(\sigma_3) \to \mathfrak{pr}(fun(\overline{\sigma}\,;\tau))$. By rule FUN it is enough to show that $\vdash^{ol} \mathfrak{pr}(\sigma_3) \le \mathfrak{pr}(\sigma_1)$ and $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_2 \le \mathfrak{pr}(fun(\overline{\sigma};\tau))$. We have the former by induction hypothesis. For the latter, by induction hypothesis we also have $\vdash^{ol} \forall \overline{a}.\rho_2 \le \mathfrak{pr}(fun(\overline{\sigma}\,;\tau))$ and by inversion this is derivable by SPEC; hence $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_2 \le \mathfrak{pr}(fun(\overline{\sigma}\,;\tau))$ and we are done.

- Case SPEC. We have $\mathfrak{pr}(\forall \overline{a}.\rho) = \sigma_1'$, $\mathfrak{pr}(fun(\overline{\sigma}\,;\tau_1)) = \sigma_2' = \rho_2'$. Moreover $\overline{\sigma} \vdash^{\partial s\ell} \forall \overline{a}.\rho \le \tau_1$, given that

$$\overline{\sigma} \vdash^{\partial s\ell} [\overline{a \mapsto \tau}]\rho \le \tau_1 \tag{4}$$

It must be that $\mathfrak{pr}(\forall \overline{a}.\rho) = \forall \overline{a}\overline{b}.\rho_1$, where $\overline{a} \notin \overline{b}$ and $\forall \overline{b}.\rho_1 = \mathfrak{pr}(\rho)$. This implies that $[\overline{a \mapsto \tau}]\forall \overline{b}.\rho_1 = \mathfrak{pr}([\overline{a \mapsto \tau}]\rho)$. From this, equation (4), and induction hypothesis $\vdash^{ol} [\overline{a \mapsto \tau}]\forall \overline{b}.\rho_1 \le \rho_2'$. Without loss of generality assume as well $\overline{b} \notin ftv(\overline{\tau})$ and then we have $\vdash^{ol} \forall \overline{b}.[\overline{a \mapsto \tau}]\rho_1 \le \rho_2'$. If $\overline{b} = \emptyset$ then we just apply SPEC and we are done. If $\overline{b} \ne \emptyset$, then by inversion it must be the case that this was derivable by SPEC, so $\vdash^{ol} [\overline{b \mapsto \tau_b}][\overline{a \mapsto \tau}]\rho_1 \le \rho_2'$ and by applying rule SPEC we get that $\vdash^{ol} \forall \overline{a}\overline{b}.\rho_1 \le \rho_2'$ as required.

- Case SKOL. Here $\mathfrak{pr}(\sigma_1) = \sigma_1'$, $\mathfrak{pr}(fun(\overline{\sigma}\,;\forall \overline{a}.\rho)) = \sigma_2'$, and $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \le \forall \overline{a}.\rho$, given that $\overline{a} \notin ftv(\overline{\sigma}, \sigma_1)$ and $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \le \rho$. Using Lemma 2.15 we get that $\mathfrak{pr}(fun(\overline{\sigma}\,;\forall \overline{a}.\rho)) = \forall \overline{a}\overline{b}.\mathfrak{pr}(fun(\overline{\sigma}\,;\rho_1))$, where $\forall \overline{b}.\rho_1 = \mathfrak{pr}(\rho)$, $\overline{b} \notin ftv(\overline{\sigma}), \overline{a}$. By induction hypothesis we have that $\vdash^{ol} \sigma_1' \le \mathfrak{pr}(fun(\overline{\sigma}\,;\rho))$, therefore $\vdash^{ol} \sigma_1' \le \forall \overline{b}.\mathfrak{pr}(fun(\overline{\sigma};\rho_1))$. If $\overline{b} = \emptyset$ then we apply SKOL and get the result. If $\overline{b} \ne \emptyset$ then by Lemma 2.4 we get that $\vdash^{ol} \sigma_1' \le \mathfrak{pr}(fun(\overline{\sigma}\,;\rho_1))$ and by applying rule SKOL we get the result.

□

Corollary 2.20 (Prenex subsumption). $\vdash^{\partial s\ell} \sigma_1 \le \sigma_2$ *iff* $\vdash^{ol} \mathfrak{pr}(\sigma_1) \le \mathfrak{pr}(\sigma_2)$.

Proof. Direct consequence of Theorem 2.18 and Theorem 2.19. □

Corollary 2.21 (Reflexivity). $\vdash^{\partial s\ell} \sigma \le \sigma$.

Proof. Directly follows by Corollary 2.20 and Lemma 2.2. □

Corollary 2.22 (Transitivity). *If* $\vdash^{\partial s\ell} \sigma_1 \le \sigma_2$ *and* $\vdash^{\partial s\ell} \sigma_2 \le \sigma_3$ *then* $\vdash^{\partial s\ell} \sigma_1 \le \sigma_3$.

Proof. Directly follows by Corollary 2.20 and Lemma 2.3. □

$$\boxed{\vdash^\eta \sigma \leq \sigma'}$$

$$\frac{\overline{b} \notin ftv(\forall \overline{a}.\sigma)}{\vdash^\eta \forall \overline{a}.\sigma \leq \forall \overline{b}.[\overline{a \mapsto \tau}]\sigma} \text{ SUB} \qquad \frac{\vdash^\eta \sigma_1 \leq \sigma_2 \quad \vdash^\eta \sigma_2 \leq \sigma_3}{\vdash^\eta \sigma_1 \leq \sigma_3} \text{ TRANS}$$

$$\frac{\vdash^\eta \sigma_3 \leq \sigma_1 \quad \vdash^\eta \sigma_2 \leq \sigma_4}{\vdash^\eta \sigma_1 \to \sigma_2 \leq \sigma_3 \to \sigma_4} \text{ FUN} \qquad \frac{\vdash^\eta \sigma_1 \leq \sigma_2}{\vdash^\eta \forall a.\sigma_1 \leq \forall a.\sigma_2} \text{ ALL}$$

$$\frac{}{\vdash^\eta \forall a.\sigma_1 \to \sigma_2 \leq (\forall a.\sigma_1) \to \forall a.\sigma_2} \text{ DISTRIB}$$

Figure 5: Predicative version of F-eta subsumption

2.3   Connection of deep skolemisation and Mitchell's relation

The predicative fragment of the F-eta subsumption is given in Figure 5. Let us start by proving some lemmas about it.

Lemma 2.23 (Reflexivity). $\vdash^\eta \sigma \leq \sigma$.

Proof. Follows directly by rule SUB. □

Lemma 2.24 (Substitution). *If* $\vdash^\eta \sigma_1 \leq \sigma_2$ *then* $\vdash^{ol} S\sigma_1 \leq S\sigma_2$, *and the new derivation has the same height.*

Proof. By induction on the height of the derivation. We proceed by case analysis on the last rule used.

- Case SUB. We have that $\vdash^\eta \forall \overline{a}.\sigma \leq \forall \overline{b}.[\overline{a \mapsto \tau}]\sigma$, given that $\overline{b} \notin ftv(\forall \overline{a}.\sigma)$. Assume without loss of generality that $\overline{a}, \overline{b} \notin vars(S)$ and then we have that $S(\forall \overline{a}.\sigma) = \forall \overline{a}.S\sigma$ and $S(\forall \overline{b}.[\overline{a \mapsto \tau}]\sigma) = \forall \overline{b}.[\overline{a \mapsto S\tau}]S\sigma$ and the result follows by rule SUB, since $\overline{b} \notin ftv(\forall \overline{a}.S\sigma)$.

- Case FUN. Follows by the induction hypotheses and rule FUN.

- Case TRANS. Follows by the induction hypotheses and rule TRANS.

- Case ALL. We have that $\vdash^\eta \forall a.\sigma_1 \leq \forall a.\sigma_2$, given that $\vdash^\eta \sigma_1 \leq \sigma_2$. Consider the substitution $S \cdot [a \mapsto c]$, where $c \notin ftv(\sigma_1, \sigma_2)$, $c \notin vars(S)$. Then, by induction hypothesis $\vdash^\eta S[a \mapsto c]\sigma_1 \leq S[a \mapsto c]\sigma_2$ and by applying rule ALL we get $\vdash^\eta \forall c.S[a \mapsto c]\sigma_1 \leq \forall c.S[a \mapsto c]\sigma_2$, or equivalently, $\vdash^\eta S(\forall a.\sigma_1) \leq S(\forall a.\sigma_2)$.

- Case DISTRIB. In this case we have that $\vdash^\eta \forall a.\sigma_1 \to \sigma_2 \leq (\forall a.\sigma_1) \to \forall a.\sigma_2$. Assume without loss of generality that $a \notin vars(S)$. Then $S(\forall a.\sigma_1 \to \sigma_2) = \forall a.S\sigma_1 \to S\sigma_2$ and the result follows by DISTRIB.

□

Lemma 2.25 (Useless quantifiers). *If* $\overline{c} \notin ftv(\sigma)$ *then* $\vdash^\eta \forall \overline{c}.\sigma \leq \sigma$ *and* $\vdash^\eta \sigma \leq \forall \overline{c}.\sigma$.

Proof. The first part follows by rule SUBS for $\overline{a} = \overline{c}$, $\overline{b} = \emptyset$. The second follows by rule SUBS for $\overline{a} = \emptyset$, $\overline{b} = \overline{c}$. □

Lemma 2.26. *If* $\vdash^\eta \sigma_1 \leq \sigma_2$ *then* $\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$.

Proof. By induction on the derivation of $\vdash^\eta \sigma_1 \leq \sigma_2$. We proceed by case analysis on the last rule used.

- Case SUB. We have that $\vdash^\eta \forall \overline{a}.\sigma \leq \forall \overline{b}.[\overline{a \mapsto \tau}]\sigma$ given that $\overline{b} \notin ftv(\forall \overline{a}.\sigma)$. To show that $\vdash^{\partial s\ell} \forall \overline{a}.\sigma \leq \forall \overline{b}.[\overline{a \mapsto \tau}]\sigma$ it is enough by Theorem 2.19, and assuming $\mathfrak{pr}(\sigma) = \forall \overline{c}.\rho$ to show that $\vdash^{ol} \forall \overline{ac}.\rho \leq \forall \overline{bc}.[\overline{a \mapsto \tau}]\rho$, where we assumed as well that $\overline{c} \notin \overline{a}, ftv(\overline{\tau}, \rho)$. Equivalently it is enough to show that $\vdash^{ol} \forall \overline{ac}.\rho \leq [\overline{a \mapsto \tau}]\rho$ or by SPEC, that $\vdash^{ol} [\overline{a \mapsto \tau}, \overline{c \mapsto c}]\rho \leq [\overline{a \mapsto \tau}]\rho$, and this follows directly by reflexivity of $\vdash^{ol}$, Lemma 2.2.

- Case TRANS. Follows by induction hypothesis and Corollary 2.22.

- Case FUN. In this case we have that $\sigma_1 = \sigma_{11} \to \sigma_{12}$ and $\sigma_2 = \sigma_{21} \to \sigma_{22}$. Moreover $\vdash^\eta \sigma_{21} \leq \sigma_{11}$ and $\vdash^\eta \sigma_{12} \leq \sigma_{22}$. By induction hypothesis

$$\vdash^{\partial s\ell} \sigma_{21} \leq \sigma_{11} \tag{1}$$

$$\vdash^{\partial s\ell} \sigma_{12} \leq \sigma_{22} \tag{2}$$

  The result then follows from (1), (2), Lemma 2.16, and an application of RFUN.

- Case ALL. We have that $\vdash^\eta \forall a.\sigma_1 \leq \forall a.\sigma_2$ given that $\vdash^\eta \sigma_1 \leq \sigma_2$. By induction hypothesis $\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$. By Theorem 2.19 $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\sigma_2)$ and assume that $\mathfrak{pr}(\sigma_1) = \forall \overline{a}_1.\rho_1, \mathfrak{pr}(\sigma_2) = \forall \overline{a}_2.\rho_2$. Equivalentlty $\vdash^{ol} \forall \overline{a}_1.\rho_1 \leq \forall \overline{a}_2.\rho_2$. By inversion $\vdash^{ol} [\overline{a_1 \to \tau}]\rho_1 \leq \rho_2$ assuming without loss of generality that $\overline{a}_2 \notin ftv(\forall \overline{a}_1.\rho_1)$. Then also $\vdash^{ol} [a \mapsto a, \overline{a_1 \mapsto \tau}]\rho_1 \leq \rho_2$ and by SPEC and SKOL we get $\vdash^{ol} \forall a\overline{a}_1.\rho_1 \leq \forall a\overline{a}_2\rho_2$. Applying Theorem 2.18 we get the result.

- Case DISTRIB. We have that

$$\vdash^\eta \forall a.\sigma_1 \to \sigma_2 \leq (\forall a.\sigma_1) \to (\forall a.\sigma_2) \tag{3}$$

  Now assume that $\mathfrak{pr}(\sigma_2) = \forall \overline{b}.\rho_2$, where without loss of generality $\overline{b} \notin a, ftv(\sigma_1)$. Then we have the following:

$$\mathfrak{pr}(\forall a.\sigma_1 \to \sigma_2) = \forall a\overline{b}.\mathfrak{pr}(\sigma_1) \to \rho_2 \tag{4}$$

$$\mathfrak{pr}((\forall a.\sigma_1) \to (\forall a.\sigma_2)) = \forall a\overline{b}.\mathfrak{pr}(\forall a.\sigma_1) \to \rho_2 \tag{5}$$

  Now we know by reflexivity, Lemma 2.2, that $\vdash^{ol} \rho_2 \leq \rho_2$. Moreover by an application of Lemma 2.2 and rule SPEC for $\vdash^{ol}$ we have that $\vdash^{ol} \mathfrak{pr}(\forall \overline{a}.\sigma_1) \leq \mathfrak{pr}(\sigma_1)$. Then by rule FUN we get that $\vdash^{ol} \mathfrak{pr}(\sigma_1) \to \rho_2 \leq \mathfrak{pr}(\forall a.\sigma_1) \to \rho_2$ and by rule SPEC $\vdash^{ol} \forall a\overline{b}.\mathfrak{pr}(\sigma_1) \to \rho_2 \leq \mathfrak{pr}(\forall a.\sigma_1) \to \rho_2$. Then we can apply SKOL and get $\vdash^{ol} \forall a\overline{b}.\mathfrak{pr}(\sigma_1) \to \rho_2 \leq \forall a\overline{b}.\mathfrak{pr}(\forall a.\sigma_1) \to \rho_2$. From this, (4), (5), and Theorem 2.18 we get that $\vdash^{\partial s\ell} \forall a.\sigma_1 \to \sigma_2 \leq (\forall a.\sigma_1) \to (\forall a.\sigma_2)$.

$\square$

Lemma 2.27. *If $\overline{a} \notin ftv(\overline{\sigma})$ then $\vdash^\eta \forall \overline{a}.fun(\overline{\sigma}\,;\rho) \leq fun(\overline{\sigma}\,;\forall \overline{a}.\rho)$.*

Proof. By induction on the size of $\overline{\sigma}$. For $\overline{\sigma} = \epsilon$ we need to show that $\vdash^\eta \forall \overline{a}.\rho \leq \forall \overline{a}.\rho$. This follows by Lemma 2.23. If $\overline{\sigma} = \sigma_1, \overline{\sigma}_1$, by Lemma 2.11 and rule DISTRIB we have

$$\vdash^\eta \forall \overline{a}.\sigma_1 \to fun(\overline{\sigma}_1\,;\rho) \leq (\forall \overline{a}.\sigma_1) \to \forall \overline{a}.fun(\overline{\sigma}_1\,;\rho) \tag{1}$$

However, by Lemma 2.25 we get that $\vdash^\eta \sigma_1 \leq \forall \overline{a}.\sigma_1$ and by induction hypothesis $\vdash^\eta \forall \overline{a}.fun(\overline{\sigma}_1\,;\rho) \leq fun(\sigma_1\,;\forall \overline{a}.\rho)$. By rule FUN we have then

$$\vdash^\eta (\forall \overline{a}.\sigma_1) \to \forall \overline{a}.fun(\overline{\sigma}_1\,;\rho) \leq \sigma_1 \to fun(\overline{\sigma}_1\,;\forall \overline{a}.\rho) \tag{2}$$

Finally combining (1) and (2) with rule TRANS we get the result.

$\square$

Lemma 2.28 (Skolemisation admissibility). *If $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\rho)$ and $\overline{a} \notin ftv(\overline{\sigma}, \sigma_1)$ then $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\forall\overline{a}.\rho)$.*

Proof. From consecutive uses of ALL and the assumptions we get $\vdash^\eta \forall\overline{a}.\sigma_1 \leq \forall\overline{a}.fun(\overline{\sigma}\,;\rho)$. By Lemma 2.25 $\vdash^\eta \sigma_1 \leq \forall\overline{a}.\sigma_1$. By rule TRANS $\vdash^\eta \sigma_1 \leq \forall\overline{a}.fun(\overline{\sigma}\,;\rho)$. By Lemma 2.27 $\vdash^\eta \forall\overline{a}.fun(\overline{\sigma}\,;\rho) \leq fun(\overline{\sigma}\,;\forall\overline{a}.\rho)$. We get the result by application of rule TRANS. $\square$

Lemma 2.29 (Specialisation admissibility). *If $\vdash^\eta [\overline{a \mapsto \tau}]\sigma_1 \leq \sigma_2$ then $\vdash^\eta \forall\overline{a}.\sigma_1 \leq \sigma_2$.*

Proof. By rule SUB we have that $\vdash^\eta \forall\overline{a}.\sigma_1 \leq [\overline{a \mapsto \tau}]\sigma_1$ and the result follows by rule TRANS. $\square$

Lemma 2.30. *If $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$ then $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\sigma_2)$.*

Proof. By induction on the derivation $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$. We proceed by case analysis on the last rule used.

- Case RFUN. Here $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \leq \sigma_2 \to \sigma_3$ given that $\overline{\sigma}, \sigma_2 \vdash^{\partial s\ell} \sigma_1 \leq \sigma_3$. By induction hypothesis $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}, \sigma_2\,;\sigma_3)$ and by definition $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\sigma_2 \to \sigma_3)$.

- Case LFUN. We have that $\sigma_3, \overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \to \sigma_2 \leq \tau$ given that $\vdash^{\partial s\ell} \sigma_3 \leq \sigma_1$ and $\overline{\sigma} \vdash^{\partial s\ell} \sigma_2 \leq \tau$. By induction hypothesis $\vdash^\eta \sigma_2 \leq fun(\overline{\sigma}\,;\tau)$ and $\vdash^\eta \sigma_2 \leq \sigma_1$. With an application of FUN we get $\vdash^\eta \sigma_1 \to \sigma_2 \leq \sigma_3 \to fun(\overline{\sigma}\,;\tau)$, or equivalently $\vdash^\eta \sigma_1 \to \sigma_2 \leq fun(\sigma_3, \overline{\sigma}\,;\tau)$ as required.

- Case SKOL. In this case $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \leq \forall\overline{a}.\rho$, given that $\overline{\sigma} \vdash^{\partial s\ell} \sigma_1 \leq \rho$ and $\overline{a} \notin ftv(\overline{\sigma}, \sigma_1)$. By induction hypothesis $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\rho)$ and by Lemma 2.28 $\vdash^\eta \sigma_1 \leq fun(\overline{\sigma}\,;\forall\overline{a}.\rho)$.

- Case SPEC. We have that $\overline{\sigma} \vdash^{\partial s\ell} \forall\overline{a}.\rho \leq \tau_1$ given that $\overline{\sigma} \vdash^{\partial s\ell} [\overline{a \mapsto \tau}]\rho \leq \tau_1$. By induction hypothesis $\vdash^\eta [\overline{a \mapsto \tau}]\rho \leq fun(\overline{\sigma}\,;\tau_1)$ and by Lemma 2.29 $\vdash^\eta \forall\overline{a}.\rho \leq fun(\overline{\sigma}\,;\tau_1)$ as required.

- Case MONO. Follows by Lemma 2.23.

$\square$

Corollary 2.31. $\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$ *iff* $\vdash^\eta \sigma_1 \leq \sigma_2$.

Proof. Follows by Lemma 2.30 and Lemma 2.26. $\square$

Corollary 2.32. $(\vdash^{ol}) \subseteq (\vdash^\eta)$ *and consequently* $(\vdash^{ol}) \subseteq (\vdash^{\partial s\ell})$.

Proof. Just notice that every rule of in Odersky-Läufer subsumption is admissible in $\vdash^\eta$. Rule FUN already exists, admissibility of MONO follows by reflexivity, admissibility of SKOL and SPEC follows from Lemma 2.28 and Lemma 2.29 respectively. $\square$

$$\boxed{\vdash^{dsk} \sigma \le \sigma'}$$

$$\dfrac{\overline{a} \notin ftv(\sigma_1) \quad \vdash^{dsk} \sigma_1 \le \rho}{\vdash^{dsk} \sigma_1 \le \sigma_2} \; \text{DEEP-SKOL} \qquad \dfrac{pr(\sigma_2) = \forall \overline{a}.\rho \qquad \vdash^{dsk} [\overline{a} \mapsto \tau]\rho_1 \le \rho_2}{\vdash^{dsk} \forall \overline{a}.\rho_1 \le \rho_2} \; \text{SPEC}$$

$$\dfrac{\vdash^{dsk} \sigma_3 \le \sigma_1 \qquad \vdash^{dsk} \sigma_2 \le \sigma_4}{\vdash^{dsk} (\sigma_1 \to \sigma_2) \le (\sigma_3 \to \sigma_4)} \; \text{FUN} \qquad \dfrac{}{\vdash^{dsk} \tau \le \tau} \; \text{MONO}$$

Figure 6: Alternative deep skolemisation subsumption

2.4 Final definition of deep skolemisation subsumption

Consider the definition of *weak prenex conversion* given below.

Definition 2.33 (Weak prenex conversion). *The function $pr(\cdot): \mathcal{S} \to \mathcal{S}$ is defined as follows:*

$$pr(\sigma) = \begin{cases} \tau & \sigma = \tau \\ \forall \overline{a}.\sigma_1 \to \rho_2 & \sigma = \sigma_1 \to \sigma_2 \wedge \sigma \ne \tau \\ & pr(\sigma_2) = \forall \overline{a}.\rho_2 \\ & \overline{a} \notin ftv(\sigma_1) \\ \forall \overline{a}\,\overline{b}.\rho_2 & \sigma = \forall \overline{a}.\rho_1 \wedge \overline{a} \ne \emptyset \\ & pr(\rho_1) = \forall \overline{b}.\rho_2 \\ & \overline{b} \notin \overline{a} \end{cases}$$

This is like normal prenex conversion but does not canonicalise the argument types in arrow types[3]. Based on this definition we give a relation that is exactly like the Odersky-Läufer subsumption but in the SKOL rule performs weak prenex conversion and skolemisation of the resulting outermost quantifiers. The relation is given in Figure 6. In this section we prove that this relation is equivalent once again with $\vdash^{dsk}$ and consequently $\vdash^{\eta}$.

Lemma 2.34 (Reflexivity). $\vdash^{dsk} \sigma \le \sigma$.

Proof. Easy induction. □

Lemma 2.35. $\vdash^{dsk} pr(\sigma) \le \sigma$ *and* $\vdash^{dsk} \sigma \le pr(\sigma)$.

Proof. By induction on σ appealing to reflexivity of $\vdash^{dsk}$, Lemma 2.34. □

Lemma 2.36. $\mathfrak{pr}(\sigma) = \forall \overline{a}.\rho_1$ *iff* $pr(\sigma) = \forall \overline{a}.\rho_2$ *and* $\mathfrak{pr}(\rho_2) = \rho_1$.

Proof. Easily follows by the definitions. □

Definition 2.37 (Canonical and prenex-canonical derivations). *Let $\mathcal{D}$ denote a derivation tree of $\vdash^{dsk}$. We say that $\mathcal{D}$ is a* canonical derivation *iff rule SKOL is used once as the very last rule, and as the very last rule of the left subtree of a tree ending with FUN. We also use the term* prenex-canonical derivation *to refer to a canonical derivation where the last rule used was a trivial application of rule SKOL and can be therefore omitted—equivalently the second type is a ρ-type in weak prenex form already.*

---

[3]Notice as well that this relation is no longer syntax-directed.

Obviously a canonical derivation can be decomposed into a prenex-canonical derivation and a non-trivial application of rule SKOL.

Lemma 2.38 (Substitution). *If $\vdash^{dsk} \sigma_1 \leq \sigma_2$ then $\vdash^{dsk} S\sigma_1 \leq S\sigma_2$, and the new derivation has the same height and the same "shape" (that is, the rules in the new derivation are applied in the same order).*

Proof. Similar to the proof of Lemma 2.1, with an easy lemma for commuting $pr(\cdot)$ and substitution. ☐

From Lemma 2.38 we get that substitution preserves canonical, or prenex-canonical derivations as a corollary.

Lemma 2.39. *If $\vdash^{dsk} \sigma_1 \leq \sigma_2$ then there exists a canonical derivation of this with the same height.*

Proof. By induction on $\vdash^{dsk} \sigma_1 \leq \sigma_2$. We proceed with case analysis on the last rule used.

- Case SKOL. In this case we have that $\vdash^{dsk} \sigma_1 \leq \sigma_2$, given that

$$pr(\sigma_2) = \forall \overline{a}.\rho \tag{1}$$

$$\overline{a} \notin ftv(\sigma_1) \tag{2}$$

$$\vdash^{dsk} \sigma_1 \leq \rho \tag{3}$$

  By induction, there exists a canonical derivation of $\vdash^{dsk} \sigma_1 \leq \rho$, in which SKOL is used above FUN and is the last rule in the derivation. But it is easy to confirm that $pr(\rho) = \rho$, therefore the last application of rule SKOL in that derivation was a trivial application and can be omitted. Then the new derivation is canonical.

- Case SPEC. Here we have that $\vdash^{dsk} \forall \overline{a}.\rho_1 \leq \rho_2$ given that

$$\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2 \tag{4}$$

  Now by induction there exists a canonical derivation of (4). The last rule is SKOL and we have that

$$pr(\rho_2) = \forall \overline{b}.\rho_2^* \tag{5}$$

$$\overline{b} \notin ftv([\overline{a \mapsto \tau}]\rho_1) \tag{6}$$

$$\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2^* \tag{7}$$

  where (7) does not have SKOL as the last rule used. Then by (7) and rule SPEC we get $\vdash^{dsk} \forall \overline{a}.\rho_1 \leq \rho_2^*$. By (6) it must also be that $\overline{b} \notin ftv(\forall \overline{a}.\rho_1)$. From this and (5) we can apply rule SKOL to get that $\vdash^{dsk} \forall \overline{a}.\rho_1 \leq \rho_2$.

- Case FUN. In this case $\vdash^{dsk} (\sigma_1 \to \sigma_2) \leq (\sigma_3 \to \sigma_4)$, given that

$$\vdash^{dsk} \sigma_3 \leq \sigma_1 \tag{8}$$

$$\vdash^{dsk} \sigma_2 \leq \sigma_4 \tag{9}$$

  By induction hypothesis there exists a canonical derivation of (8) with the same height. Moreover by induction there exists a canonical derivation of (9) with the same height. The last rule used was SKOL and we have that

$$pr(\sigma_4) = \forall \overline{a}.\rho_4 \tag{10}$$

$$\overline{a} \notin ftv(\sigma_2) \tag{11}$$

$$\vdash^{dsk} \sigma_2 \leq \rho_4 \tag{12}$$

  Assume without loss of generality that $\overline{a} \notin ftv(\sigma_1, \sigma_3, \sigma_2)$ as well. Then notice that $pr(\sigma_3 \to \sigma_4) = \forall \overline{a}.\sigma_3 \to \rho_4$. By FUN and (8) and (12) we get that $\vdash^{dsk} \sigma_1 \to \sigma_2 \leq \sigma_3 \to \rho_4$ and by applying rule SKOL we get the result.

- Case MONO. Just apply trivially rule SKOL as the last rule of the derivation.

$\square$

The above lemma shows that there is an algorithmic implementation that applies deep skolemisation at the begining and when comparing the argument types of two functions only. The corresponding syntax-directed presentation can be found in the main paper. In this document we use the non syntax-directed presentation in combination with the the canonical and prenex-canonical derivations lemma.

Lemma 2.40. *If* $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\sigma_2)$ *then* $\vdash^{dsk} \sigma_1 \leq \sigma_2$.

Proof. By induction on the lexicographic pair of the height of the derivation of $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\sigma_2)$ and the number of recursive calls of $\mathfrak{pr}(\sigma_1)$. We have the following cases to consider.

- Case SKOL. Here $\mathfrak{pr}(\sigma_1) = \sigma$ and $\mathfrak{pr}(\sigma_2) = \forall \overline{a}.\rho$. Then consider $pr(\sigma_2) = \forall \overline{a}.\rho'$. It must be that $\rho = \mathfrak{pr}(\rho')$ by Lemma 2.36. Moreover $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \forall \overline{a}.\mathfrak{pr}(\rho')$, given that $\overline{a} \notin ftv(\mathfrak{pr}(\sigma_1)) = ftv(\sigma)$ and

$$\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\rho') \tag{1}$$

By (1) and induction hypothesis it must be that $\vdash^{dsk} \sigma_1 \leq \rho'$ and by applying rule SKOL we get $\vdash^{dsk} \sigma_1 \leq \sigma_2$, since $\overline{a} \notin ftv(\sigma_1)$.

- Case SPEC. For this case we have

$$\mathfrak{pr}(\sigma_1) = \forall \overline{a}.\rho_1 \tag{2}$$
$$\mathfrak{pr}(\sigma_2) = \rho_2 \tag{3}$$
$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2 \tag{4}$$

We wish to show that $\vdash^{dsk} \sigma_1 \leq \sigma_2$. We proceed with a case analysis on the structure of σ_1, and without loss of generality let us assume that σ_1 is not a monotype.

  - Case $\sigma_1 = \sigma_{11} \rightarrow \sigma_{12}$. Then $\mathfrak{pr}(\sigma_1) = \forall \overline{a}.\mathfrak{pr}(\sigma_{11}) \rightarrow \rho_{12}$ where

$$\mathfrak{pr}(\sigma_{12}) = \forall \overline{a}.\rho_{12} \tag{5}$$

and $\overline{a} \notin ftv(\sigma_{11})$. By (4) it must be that $\vdash^{ol} [\overline{a \mapsto \tau}](\mathfrak{pr}(\sigma_{11}) \rightarrow \rho_{12}) \leq \rho_2$, or $\vdash^{ol} \mathfrak{pr}(\sigma_{11}) \rightarrow [\overline{a \mapsto \tau}]\rho_{12} \leq \rho_2$. By easy inversion it must be that $\rho_2 = \sigma_2^A \rightarrow \sigma_2^B$ and

$$\vdash^{ol} \sigma_2^A \leq \mathfrak{pr}(\sigma_{11}) \tag{6}$$
$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_{12} \leq \sigma_2^B \tag{7}$$

By (3) then it must be that

$$\mathfrak{pr}(\sigma_2)) = \sigma_2^A \rightarrow \sigma_2^B \tag{8}$$

Then it must be the case that $\sigma_2 = \sigma_{21} \rightarrow \sigma_{22}$, such that $\mathfrak{pr}(\sigma_{21}) = \sigma_2^A$ and $\mathfrak{pr}(\sigma_{22}) = \sigma_2^B$. By (6) we have that $\mathfrak{pr}(\sigma_{21}) \leq \mathfrak{pr}(\sigma_{11}$ and by the induction hypothesis

$$\vdash^{dsk} \sigma_{21} \leq \sigma_{11} \tag{9}$$

Similarly by (7) we have that $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_{12} \leq \mathfrak{pr}(\sigma_{22})$ and by an application of SPEC we have $\vdash^{ol} \forall \overline{a}.\rho_{12} \leq \mathfrak{pr}(\sigma_{22})$. But the height of this derivation is still one less than the height of the derivation we are examining and therefore by induction hypothesis we have that

$$\vdash^{dsk} \sigma_{12} \leq \sigma_{22} \tag{10}$$

From (9) and (10) and FUN we get that $\vdash^{dsk} \sigma_{11} \rightarrow \sigma_{12} \leq \sigma_{21} \rightarrow \sigma_{22}$.

– Case $\sigma_1 = \forall \overline{a}_1 . \rho_{11}$. Here we have that $\mathfrak{pr}(\sigma_1) = \forall \overline{a}_1 \overline{a}_2 . \rho_{12}$, such that $\forall \overline{a}_2 . \rho_{12} = \mathfrak{pr}(\rho_{11})$. From this it follows that

$$\forall \overline{a}_2 . [\overline{a_1 \mapsto \tau_1}]\rho_{12} = \mathfrak{pr}([\overline{a_1 \mapsto \tau_1}]\rho_{11}) \tag{11}$$

From Lemma 2.17 we know that the number of recursive calls of $\mathfrak{pr}([\overline{a_1 \to \tau_1}]\rho_{11})$ is the same as of $\mathfrak{pr}(\rho_{11})$. We also know that

$$\vdash^{ol} [\overline{a_2 \mapsto \tau_2}][\overline{a_1 \mapsto \tau_1}]\rho_{12} \leq \rho_2 \tag{12}$$

Then, we can apply the induction hypothesis to get that $\vdash^{dsk} [\overline{a_1 \mapsto \tau_1}]\rho_{11} \leq \rho_2$ and by applying rule SPEC $\vdash^{dsk} \forall \overline{a}_1 . \rho_{11} \leq \rho_2$, as required.

- Case FUN. In this case we have the following

$$\mathfrak{pr}(\sigma_1) = \sigma_1' \to \sigma_2' \tag{13}$$
$$\mathfrak{pr}(\sigma_2) = \sigma_3' \to \sigma_4' \tag{14}$$
$$\vdash^{ol} \sigma_1' \to \sigma_2' \leq \sigma_3' \to \sigma_4' \tag{15}$$
$$\vdash^{ol} \sigma_3' \leq \sigma_1' \tag{16}$$
$$\vdash^{ol} \sigma_2' \leq \sigma_4' \tag{17}$$

By (13) it must be that

$$\sigma_1 = \sigma_{11} \to \sigma_{12} \tag{18}$$
$$\mathfrak{pr}(\sigma_{11}) = \sigma_1' \tag{19}$$
$$\mathfrak{pr}(\sigma_{12}) = \sigma_2' \tag{20}$$

And similarly by (14) we get

$$\sigma_2 = \sigma_{21} \to \sigma_{22} \tag{21}$$
$$\mathfrak{pr}(\sigma_{21}) = \sigma_3' \tag{22}$$
$$\mathfrak{pr}(\sigma_{22}) = \sigma_4' \tag{23}$$

By (19), (22), (16), and induction hypothesis $\vdash^{dsk} \sigma_{21} \leq \sigma_{11}$. Similarly by (20), (23), (17), and induction hypothesis $\vdash^{dsk} \sigma_{12} \leq \sigma_{22}$. We get the result by applying rule FUN.

- Case MONO. Trivially follows by definition of $\mathfrak{pr}(\cdot)$ and rule MONO.

$\square$

Lemma 2.41. *If* $\vdash^{dsk} \sigma_1 \leq \sigma_2$ *then* $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\sigma_2)$.

Proof. By induction on the derivation $\vdash^{dsk} \sigma_1 \leq \sigma_2$. We proceed by case analysis on the last rule used.

- Case MONO. Directly folows by rule MONO.

- Case FUN. In this case we have that

$$\vdash^{dsk} \sigma_1 \to \sigma_2 \leq \sigma_3 \to \sigma_4 \tag{1}$$
$$\vdash^{dsk} \sigma_3 \leq \sigma_1 \tag{2}$$
$$\vdash^{dsk} \sigma_2 \leq \sigma_4 \tag{3}$$

Moreover we know that $\mathfrak{pr}(\sigma_1 \to \sigma_2) = \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2$, with $\forall \overline{a}.\rho_2 = \mathfrak{pr}(\sigma_2)$ and $\overline{a} \notin ftv(\sigma_1)$. On the other hand $\mathfrak{pr}(\sigma_3 \to \sigma_4) = \forall \overline{b}.\mathfrak{pr}(\sigma_3) \to \rho_4$ where $\mathfrak{pr}(\sigma_4) = \forall \overline{b}.\rho_4$ and assume that $\overline{b} \notin ftv(\sigma_3, \sigma_2, \sigma_1)$. We need to show that $\vdash^{ol} \forall \overline{a}.\mathfrak{pr}(\sigma_1) \to \rho_2 \leq \forall \overline{b}.\mathfrak{pr}(\sigma_3) \to \rho_4$. By rule SPEC it is enough to show that $\vdash^{ol} \mathfrak{pr}(\sigma_1) \to [\overline{a \mapsto \tau}]\rho_2 \leq \mathfrak{pr}(\sigma_3) \to \rho_4$. By rule SKOL and FUN it is enough to show that $\vdash^{ol} \mathfrak{pr}(\sigma_3) \leq \mathfrak{pr}(\sigma_1)$ and $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_2 \leq \rho_4$. We have the former by induction hypothesis. For the latter, by induction hypothesis on (3) we also have $\vdash^{ol} \forall \overline{a}.\rho_2 \leq \forall \overline{b}.\rho_4$ and by inversion this is derivable by SPEC; hence $\vdash^{ol} [\overline{a \mapsto \tau}]\rho_2 \leq \rho_4$ and we are done.

- Case SPEC. We have $\vdash^{dsk} \forall \overline{a}.\rho_1 \leq \rho_2$ given that $\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2$. By induction hypothesis $\vdash^{ol} \mathfrak{pr}([\overline{a \mapsto \tau}]\rho_1) \leq \mathfrak{pr}(\rho_2)$, or $\vdash^{ol} [\overline{a \mapsto \tau}]\mathfrak{pr}(\rho_1) \leq \mathfrak{pr}(\rho_2)$, or $\vdash^{ol} \forall \overline{a}.\mathfrak{pr}(\rho_1) \leq \mathfrak{pr}(\rho_2)$ by transitivity of $\vdash^{ol}$. Equivalently, by definition of $\mathfrak{pr}(\cdot)$, $\vdash^{ol} \mathfrak{pr}(\forall \overline{a}.\rho_1) \leq \mathfrak{pr}(\rho_2)$ as required.

- Case SKOL. Here $\vdash^{dsk} \sigma_1 \leq \sigma_2$ given that

$$pr(\sigma_2) = \forall \overline{a}.\rho \tag{4}$$

$$\overline{a} \notin ftv(\sigma_1) \tag{5}$$

$$\vdash^{dsk} \sigma_1 \leq \rho \tag{6}$$

It must be that $\mathfrak{pr}(\sigma_2) = \forall \overline{a}.\rho'$ such that $\mathfrak{pr}(\rho) = \rho'$. Then by (6) and induction hypothesis $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \rho'$ and by an application of rule SKOL with (5), we get $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \forall \overline{a}.\rho'$, or equivalently $\vdash^{ol} \mathfrak{pr}(\sigma_1) \leq \mathfrak{pr}(\sigma_2)$ as required.

$\square$

Corollary 2.42. $\vdash^{\mathfrak{dsk}} \sigma_1 \leq \sigma_2$ *iff* $\vdash^{\eta} \sigma_1 \leq \sigma_2$ *iff* $\vdash^{dsk} \sigma_1 \leq \sigma_2$.

Proof. Follows by Lemma 2.40, Lemma 2.41, Corollary 2.20, and Corollary 2.31. $\square$

Having Corollary 2.42 enables us to switch between all the different formalisations of Mitchell's F-eta subsumption viewing all inference rules as theorems that hold independently of the formalisation we use each time.

3 Higher-rank type systems

In this section we study type systems that support higher-rank types. We assume that the type systems of Figure 7, Figure 8, and Figure 9 rely on a **reflexive, transitive relation $\vdash^{sub\sigma}$ for which the substitution lemma holds**. This relation will stand either for deep skolemisation subsumption $\vdash^{\mathfrak{dsk}}$ or the original Odersky-Läufer subsumption relation $\vdash^{ol}$.

Several properties of the type systems hold independently of whether $\vdash^{sub\sigma}$ is $\vdash^{\mathfrak{dsk}}$ or $\vdash^{ol}$. For the lemmas that are sensitive on the exact definition of $\vdash^{sub\sigma}$ we explicitly specify what $\vdash^{sub\sigma}$ is.

In the following, we use the syntax:

$$\overline{\Gamma}(\rho) = \forall \overline{a}.\rho \text{ where } \overline{a} = ftv(\rho) - ftv(\Gamma)$$

$$\text{Rho-types} \quad \rho \ ::= \ \tau \mid \sigma \to \sigma'$$

$$\boxed{\Gamma \vdash t : \sigma}$$

$$\frac{}{\Gamma \vdash i : \texttt{Int}} \ \text{INT} \qquad \frac{}{\Gamma, (x : \sigma) \vdash x : \sigma} \ \text{VAR}$$

$$\frac{\Gamma, x : \tau \vdash t : \sigma}{\Gamma \vdash (\backslash x\,.\,t) : (\tau \to \sigma)} \ \text{ABS} \qquad \frac{\Gamma, x : \sigma \vdash t : \sigma'}{\Gamma \vdash (\backslash(x::\sigma)\,.\,t) : (\sigma \to \sigma')} \ \text{AABS}$$

$$\frac{\begin{array}{c} \Gamma \vdash t : (\sigma \to \sigma') \\ \Gamma \vdash u : \sigma \end{array}}{\Gamma \vdash t \ u : \sigma'} \ \text{APP} \qquad \frac{\begin{array}{c} \Gamma \vdash u : \sigma \\ \Gamma, x : \sigma \vdash t : \sigma' \end{array}}{\Gamma \vdash \texttt{let} \ x \ = \ u \ \texttt{in} \ t : \sigma'} \ \text{LET}$$

$$\frac{\Gamma \vdash t : \sigma}{\Gamma \vdash (t::\sigma) : \sigma} \ \text{ANNOT}$$

$$\frac{\begin{array}{c} \overline{a} \notin ftv(\Gamma) \\ \Gamma \vdash t : \rho \end{array}}{\Gamma \vdash t : \forall \overline{a}.\rho} \ \text{GEN} \qquad \frac{\begin{array}{c} \Gamma \vdash t : \sigma' \\ \vdash^{sub\sigma} \sigma' \le \sigma \end{array}}{\Gamma \vdash t : \sigma} \ \text{SUBS}$$
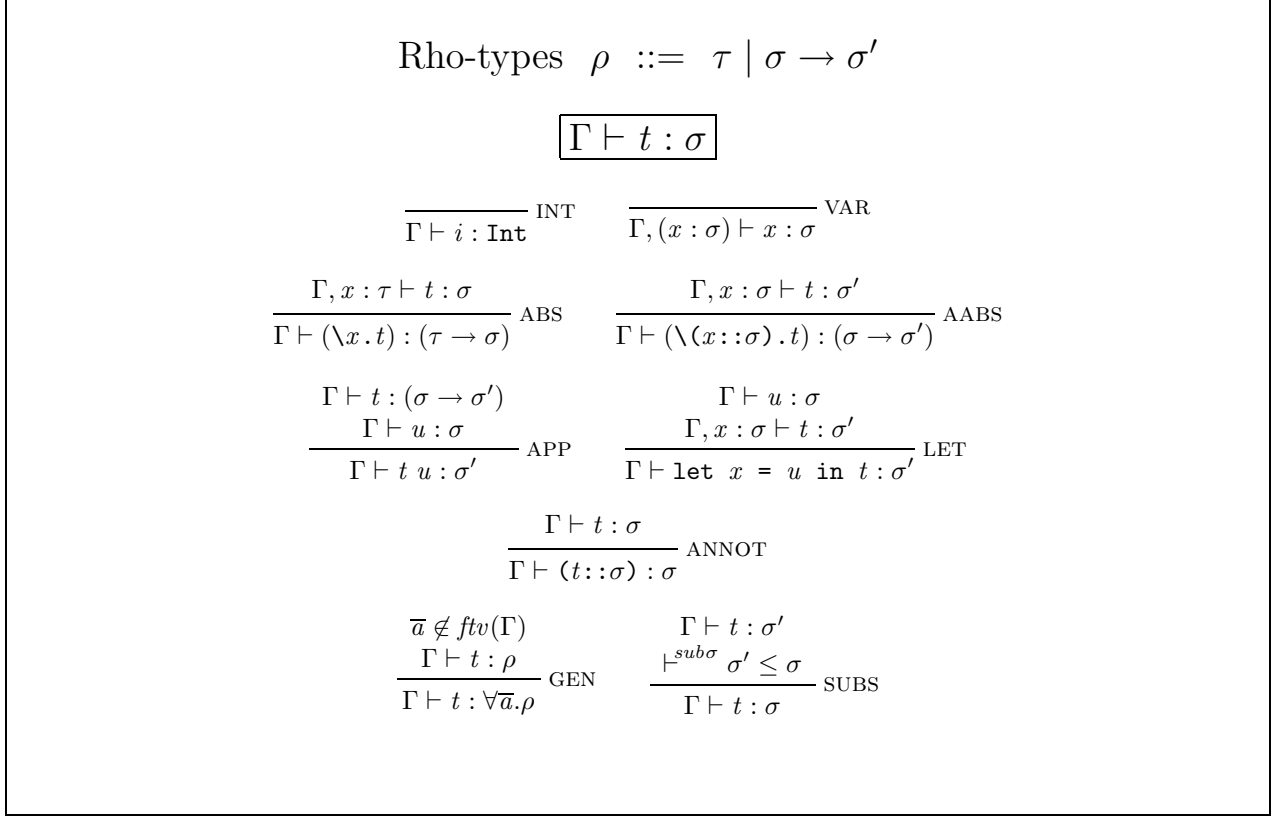
Figure 7: Non syntax-directed higher-rank type system

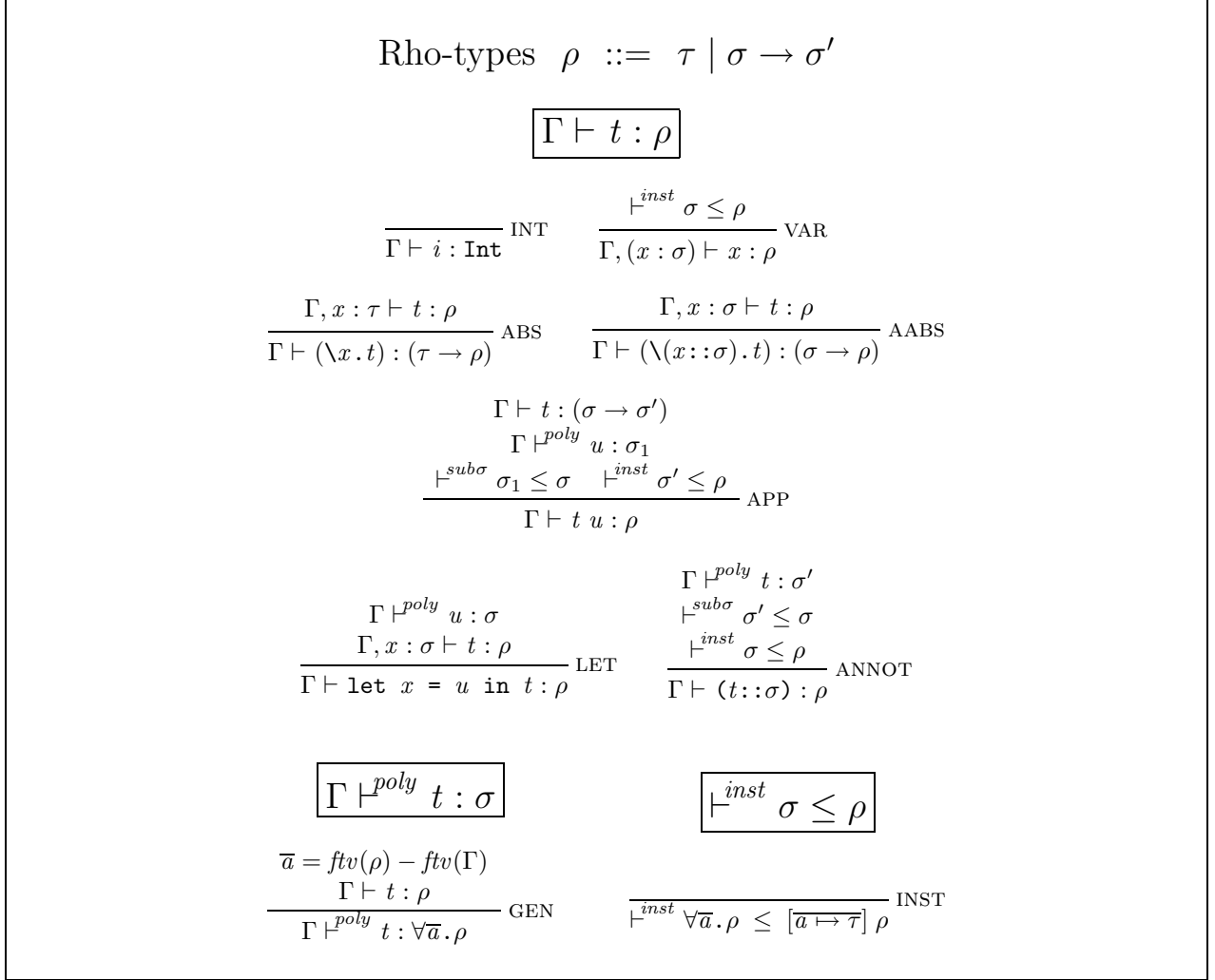## 3.1 Syntax-directed higher-rank type system

In this section we explore the connection between Figure 7 and Figure 8. For clarity let us refer to the typing relation of the non syntax-directed system of Figure 7 as $\vdash_{nsd}$, and to the typing relations of the syntax-directed version of Figure 8 as $\vdash_{sd}$ and $\vdash_{sd}^{poly}$.

Lemma 3.1. *Let $\vdash^{sub\sigma}$ be either $\vdash^{ol}$ or $\vdash^{\mathfrak{dst}}$. Then*

1. *if $\Gamma \vdash_{sd} t : \rho$ then $\Gamma \vdash_{nsd} t : \rho$.*

2. *if $\Gamma \vdash_{sd}^{poly} t : \sigma$ then $\Gamma \vdash_{nsd} t : \sigma$.*

Proof. We prove the two claims simultaneously by induction on the height of the syntax-directed derivation. We proceed by case analysis on the last rule used in the derivation. For the first part we have the cases below.

- Case INT. Directly follows by rule INT.

- Case VAR. We have that $\Gamma \vdash_{sd} x : \rho$ given that $x : \sigma \in \Gamma$ and $\vdash^{inst} \sigma \le \rho$. By rule VAR $\Gamma \vdash_{nsd} x : \sigma$. Moreover it is easy to see that $\vdash^{sub\sigma} \sigma \le \rho$. By SUBS then $\Gamma \vdash_{nsd} x : \rho$.

- Case ABS. Follows by induction hypothesis and rule ABS.

- Case AABS. Follows by induction hypothesis and rule AABS.

$$\text{Rho-types} \quad \rho \ ::= \ \tau \mid \sigma \to \sigma'$$

$$\boxed{\Gamma \vdash t : \rho}$$

$$\frac{}{\Gamma \vdash i : \texttt{Int}} \ \text{INT} \qquad \frac{\vdash^{inst} \sigma \leq \rho}{\Gamma, (x : \sigma) \vdash x : \rho} \ \text{VAR}$$

$$\frac{\Gamma, x : \tau \vdash t : \rho}{\Gamma \vdash (\backslash x . t) : (\tau \to \rho)} \ \text{ABS} \qquad \frac{\Gamma, x : \sigma \vdash t : \rho}{\Gamma \vdash (\backslash (x :: \sigma) . t) : (\sigma \to \rho)} \ \text{AABS}$$

$$\frac{\begin{array}{c} \Gamma \vdash t : (\sigma \to \sigma') \\ \Gamma \vdash^{poly} u : \sigma_1 \\ \vdash^{sub\sigma} \sigma_1 \leq \sigma \quad \vdash^{inst} \sigma' \leq \rho \end{array}}{\Gamma \vdash t \ u : \rho} \ \text{APP}$$

$$\frac{\begin{array}{c} \Gamma \vdash^{poly} u : \sigma \\ \Gamma, x : \sigma \vdash t : \rho \end{array}}{\Gamma \vdash \texttt{let} \ x \ = \ u \ \texttt{in} \ t : \rho} \ \text{LET} \qquad \frac{\begin{array}{c} \Gamma \vdash^{poly} t : \sigma' \\ \vdash^{sub\sigma} \sigma' \leq \sigma \\ \vdash^{inst} \sigma \leq \rho \end{array}}{\Gamma \vdash (t :: \sigma) : \rho} \ \text{ANNOT}$$

$$\boxed{\Gamma \vdash^{poly} t : \sigma} \qquad\qquad \boxed{\vdash^{inst} \sigma \leq \rho}$$

$$\frac{\begin{array}{c} \overline{a} = ftv(\rho) - ftv(\Gamma) \\ \Gamma \vdash t : \rho \end{array}}{\Gamma \vdash^{poly} t : \forall \overline{a} . \rho} \ \text{GEN} \qquad \frac{}{\vdash^{inst} \forall \overline{a} . \rho \ \leq \ [\overline{a \mapsto \tau}] \, \rho} \ \text{INST}$$

Figure 8: Syntax-directed higher-rank type system

- Case APP. We have that $\Gamma \vdash_{sd} t \ u : \rho$, given that

$$\Gamma \vdash_{sd} t : (\sigma \to \sigma') \tag{1}$$

$$\Gamma \vdash^{poly}_{sd} u : \sigma_1 \tag{2}$$

$$\vdash^{sub\sigma} \sigma_1 \leq \sigma \tag{3}$$

$$\vdash^{inst} \sigma' \leq \rho \tag{4}$$

By induction hypothesis and (1) we get $\Gamma \vdash_{nsd} t : \sigma \to \sigma'$. Moreover from (3) we see that $\vdash^{sub\sigma} \sigma \to \sigma' \leq \sigma_1 \to \sigma'$ because rule FUN is admissible in both $\vdash^{ol}$ and $\vdash^{\mathfrak{dsl}}$. Then by rule SUBS $\Gamma \vdash_{nsd} t : \sigma_1 \to \sigma'$. Finally by induction hypothesis for (2) we have $\Gamma \vdash_{nsd} u : \sigma_1$. Applying rule APP gives $\Gamma \vdash_{nsd} t \ u : \sigma'$ and since $(\vdash^{inst}) \subseteq (\vdash^{sub\sigma})$ we can apply rule SUBS to get $\Gamma \vdash_{nsd} t \ u : \rho$.

- Case LET. Follows by induction hypothesis and rule LET.

- Case ANNOT. In this case $\Gamma \vdash (t::\sigma) : \rho$ given that

$$\Gamma \vdash_{sd}^{poly} t : \sigma' \tag{5}$$

$$\vdash^{sub\sigma} \sigma' \leq \sigma \tag{6}$$

$$\vdash^{inst} \sigma \leq \rho \tag{7}$$

By (5) and induction hypothesis $\Gamma \vdash_{nsd} t : \sigma'$. By rule SUBS $\Gamma \vdash_{nsd} t : \sigma$ and moreover from (6) we have $\vdash^{inst} \sigma \leq \rho$. Applying rule SUBS once again gives $\Gamma \vdash_{nsd} t : \rho$ as required.

For the second part we have that $\Gamma \vdash_{sd}^{poly} t : \forall \overline{a}.\rho$ when $\overline{a} = ftv(\rho) - ftv(\Gamma)$ and $\Gamma \vdash_{sd} t : \rho$. By induction hypothesis $\Gamma \vdash_{nsd} t : \rho$ and by rule GEN $\Gamma \vdash_{nsd} t : \forall \overline{a}.\rho$. $\qquad\square$

Lemma 3.2 (Substitution).

1. *If* $\Gamma \vdash_{sd} t : \rho$ *then* $S\Gamma \vdash_{sd} t : S\rho$.

2. *If* $\Gamma \vdash_{sd}^{poly} t : \sigma$ *then* $S\Gamma \vdash_{sd}^{poly} t : S\sigma$.

Proof. Straightforward induction appealing to the substitution property for $\vdash^{sub\sigma}$. $\qquad\square$

Lemma 3.3 (Weakening). *Assume that* $\vdash^{sub\sigma}$ *is* $\vdash^{\partial s\ell}$. *If* $\Gamma_2 \vdash_{sd} t : \rho_2$ *and* $\vdash^{\partial s\ell} \Gamma_1 \leq \Gamma_2$ *then* $\Gamma_1 \vdash_{sd} t : \rho_1$ *with* $\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_1) \leq \overline{\Gamma_2}(\rho_2)$. *Consequently if* $\Gamma_2 \vdash_{sd}^{poly} t : \sigma_2$ *then* $\Gamma_1 \vdash_{sd}^{poly} t : \sigma_1$ *with* $\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$.

Proof. The proof is by induction on the height of the derivation. We proceed with case analysis on the last rule used in the derivation.

- Case INT. Directly follows by rule INT.
- Case VAR. We have that $\Gamma_2 \vdash_{sd} x : \rho_2$ given that $x : \sigma_2 \in \Gamma_2$ and

$$\vdash^{inst} \sigma_2 \leq \rho_2 \tag{1}$$

Then, $x : \sigma_1 \in \Gamma_1$ such that

$$\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2 \tag{2}$$

Assume that $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2)$. Then by the substitution lemma we get $\Gamma_2 \vdash_{sd} x : [\overline{b \mapsto d}]\rho_2$ for some $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \rho_2)$. By (1) and the substitution lemma we get $\vdash^{inst} \sigma_2 \leq [\overline{b \mapsto d}]\rho_2$ and by transitivity of $\vdash^{\partial s\ell}$

$$\vdash^{\partial s\ell} \sigma_1 \leq [\overline{b \mapsto d}]\rho_2 \tag{3}$$

Moreover assume that $\sigma_1 = \forall \overline{a}.\rho_1$ and without loss of generality assume that $\overline{a} \notin ftv(\Gamma_1)$. Then $\vdash^{inst} \sigma_1 \leq \rho_1$. By (3) we get that $\vdash^{\partial s\ell} \forall \overline{a}.\rho_1 \leq [\overline{b \mapsto d}]\rho_2$. Consider $\overline{a}' = ftv(\rho_1) - ftv(\Gamma_1)$. Then $\overline{a} \subseteq \overline{a}'$ and consequently $\vdash^{\partial s\ell} \forall \overline{a}'.\rho_1 \leq [\overline{b \mapsto d}]\rho_2$. Then it must be that $\overline{d} \notin ftv(\forall \overline{a}'.\rho_1)$ because otherwise $\overline{d} \in ftv(\Gamma_1)$. Then by SKOL admissibility $\vdash^{\partial s\ell} \forall \overline{a}'.\rho_1 \leq \forall \overline{d}.[\overline{b \mapsto d}]\rho_2$, or equivalently $\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_1) \leq \overline{\Gamma_2}(\rho_2)$.

- Case ABS. Here we have that $\Gamma_2 \vdash_{sd} (\backslash x . t) : (\tau \to \rho_2)$, given that

$$\Gamma_2, x : \tau \vdash_{sd} t : \rho_2 \tag{4}$$

Consider $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2, \tau)$ and a renaming substitution $[\overline{b \mapsto d}]$ where $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \tau, \rho_2)$. Then by (4) and the substitution lemma we get $\Gamma_2, x : \tau \vdash_{sd} t : [\overline{b \mapsto d}]\rho_2$. By induction hypothesis there exists a ρ_1 such that $\Gamma_1, x : \tau \vdash_{sd} t : \rho_1$ and

$$\vdash^{\partial s\ell} \forall \overline{a} . \rho_1 \leq \forall \overline{d} . [\overline{b \mapsto d}]\rho_2 \tag{5}$$

where $\overline{a} = ftv(\rho_1) - ftv(\Gamma_1, \tau)$. By rule ABS we get $\Gamma_1 \vdash_{sd} (\backslash x . t) : (\tau \to \rho_1)$. We wish to show that

$$\vdash^{\partial s\ell} \forall \overline{a}_1 . \tau \to \rho_1 \leq \forall \overline{a}_2 . \tau \to \rho_2 \tag{6}$$

where $\overline{a}_1 = ftv(\rho_1, \tau) - ftv(\Gamma)$ and $\overline{a}_2 = ftv(\rho_2, \tau) - ftv(\Gamma)$. Notice that if $\overline{c} = ftv(\tau) - ftv(\Gamma)$ then $\overline{a}_1 = \overline{ac}$ and $\overline{a}_2 = \overline{bc}$. From (5), since by SUB $\vdash^{\partial s\ell} \forall \overline{d} . [\overline{b \mapsto d}]\rho_2 \leq [\overline{b \mapsto d}]\rho_2$, and by transitivity we get that

$$\vdash^{\partial s\ell} \forall \overline{a} . \rho_1 \leq [\overline{b \mapsto d}]\rho_2 \tag{7}$$

Then, by rule FUN $\vdash^{\partial s\ell} \tau \to \forall \overline{a} . \rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$ and $\vdash^{\partial s\ell} \forall \overline{a} . \tau \to \rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$ by transitivity and rule DISTRIB. By SUB and transitivity $\vdash^{\partial s\ell} \forall \overline{ac} . \tau \to \rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$. Now we claim that $\overline{c} \notin ftv(\forall \overline{ac} . \tau \to \rho_1)$ and $\overline{d} \notin ftv(\forall \overline{ac} . \tau \to \rho_1)$ similarly. The former because we quantified over them, the latter because the opposite would mean that $\overline{d} \in ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get that $\vdash^{\partial s\ell} \forall \overline{ac} . \tau \to \rho_1 \leq \forall \overline{dc} . \tau \to [\overline{b \mapsto d}]\rho_2$ and by an α-renaming of $\overline{d}$ to $\overline{b}$ we are done.

- Case AABS. Similar to the case for ABS.

- Case APP. Here we have $\Gamma_2 \vdash_{sd} t\, u : \rho_2$ given that

$$\Gamma_2 \vdash_{sd} t : (\sigma \to \sigma') \tag{8}$$
$$\Gamma_2 \vdash_{sd} u : \sigma_1 \tag{9}$$
$$\vdash^{\partial s\ell} \sigma_1 \leq \sigma \tag{10}$$
$$\vdash^{inst}_{\Uparrow} \sigma' \leq \rho_2 \tag{11}$$

Consider a renaming substitution $\overline{g}_1 = ftv(\sigma, \sigma', \rho_2) - ftv(\Gamma_2)$ to fresh $\overline{g}_2$, such that $\overline{g}_2 \notin ftv(\Gamma_1)$. Then by the substitution lemma (8) becomes

$$\Gamma_2 \vdash_{sd} t : (\sigma_0 \to \sigma'_0) \tag{12}$$

where $\sigma_0 = [\overline{g_1 \mapsto g_2}]\sigma$ and $\sigma'_0 = [\overline{g_1 \mapsto g_2}]\sigma'$. By induction hypothesis on (12) we get that there exists a ρ_t with

$$\Gamma_1 \vdash_{sd} t : \rho_t \tag{13}$$
$$\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_t) \leq \forall \overline{g}_2 . \sigma_0 \to \sigma'_0 \tag{14}$$

Because of the choice of $\overline{g}_2$, from equation (14) we get

$$\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_t) \leq \sigma_0 \to \sigma'_0 \tag{15}$$

There are two possible cases for ρ_t. It is either a type variable $a \notin ftv(\Gamma_1)$ or it will be an arrow type $\sigma^t_1 \to \sigma^t_2$.

26

– Assume that $\rho_t = \sigma_1^t \to \sigma_2^t$ and let $\overline{a} = ftv(\rho_t) - ftv(\Gamma_1)$. Then by equation (15) and Corollary 2.20 we get:

$$\vdash^{ol} \forall \overline{a}\overline{b} . \mathfrak{pr}(\sigma_1^t) \to \rho_2^t \le \forall \overline{c} . \mathfrak{pr}(\sigma_0) \to \rho_0' \tag{16}$$

where

$$\overline{b} \notin ftv(\sigma_1^t) \tag{17}$$
$$\overline{c} \notin ftv(\sigma_0, \Gamma_1, \Gamma_2, \sigma_1^t, \sigma_2^t) \tag{18}$$
$$\mathfrak{pr}(\sigma_2^t) = \forall \overline{b} . \rho_2^t \tag{19}$$
$$\mathfrak{pr}(\sigma_0') = \forall \overline{c} . \rho_0' \tag{20}$$

By (16) and (18) it must be that

$$
\begin{aligned}
& \vdash^{ol} \forall \overline{a}\overline{b} . \mathfrak{pr}(\sigma_1^t) \to \rho_2^t \le \mathfrak{pr}(\sigma_0) \to \rho_0' && \text{(by inversion)} \\
\Rightarrow \quad & \vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}](\mathfrak{pr}(\sigma_1^t) \to \rho_2^t) \le \mathfrak{pr}(\sigma_0) \to \rho_0' && \text{(by inversion)} \\
\Rightarrow \quad & \vdash^{ol} [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \to [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \mathfrak{pr}(\sigma_0) \to \rho_0'
\end{aligned}
$$

From the last equation, by inversion we get that

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \le [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \tag{21}$$
$$\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \rho_0' \tag{22}$$

From Corollary 2.20 and (21)

$$\vdash^{\partial s \ell} \sigma_0 \le [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{23}$$

From (13) and the substitution lemma, we get

$$\Gamma_1 \vdash_{sd} t : [\overline{a \mapsto \tau_a}](\sigma_1^t \to \sigma_2^t) \tag{24}$$

By the substitution lemma for (9) we have that $\Gamma_2 \vdash_{sd}^{poly} u : [\overline{g_1 \mapsto g_2}]\sigma_1$ and by transitivity of $\vdash^{\partial s \ell}$ we get:

$$\vdash^{\partial s \ell} [\overline{g_1 \mapsto g_2}]\sigma_1 \le [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{25}$$

Moreover, by induction hypothesis we have that $\Gamma_1 \vdash_{sd}^{poly} u : \sigma_1'$ such that $\vdash^{\partial s \ell} \sigma_1' \le [\overline{g_1 \mapsto g_2}]\sigma_1$. From this and (25) we get $\vdash^{\partial s \ell} \sigma_1' \le [\overline{a \mapsto \tau_a}]\sigma_1^t$. Then, if $[\overline{a \mapsto \tau_a}]\sigma_2^t = \forall \overline{g_3} . \rho_1$, where without loss of generality $\overline{g_3} \notin ftv(\Gamma_1)$ we have that $\vdash^{inst} [\overline{a \mapsto \tau_a}]\sigma_2^t \le \rho_1$. We have all the premises of the rule APP and applying it gives us that $\Gamma_1 \vdash t\ u : \rho_1$. Then it is the case that $\overline{\Gamma_1}(\rho_1) = \overline{\Gamma_1}([\overline{a \mapsto \tau_a}]\sigma_2^t)$. By (22) we get that

$$
\begin{aligned}
& \vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \rho_0' \\
\Rightarrow \quad & \vdash^{ol} \forall \overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \rho_0' \\
\Rightarrow \quad & \vdash^{ol} \forall \overline{d}\overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \rho_0'
\end{aligned}
$$

where $\overline{d} = ftv(\overline{\tau}_a, \rho_2^t) - ftv(\Gamma_1)$. But now we know that $\overline{c} \notin ftv(\forall \overline{d}\overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t)$, because it must be that $ftv(\forall \overline{d}\overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t) \subseteq ftv(\Gamma_1)$, and by (18) $\overline{c} \notin ftv(\Gamma_1)$. Then we can apply rule SKOL to get that $\vdash^{ol} \forall \overline{d}\overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \forall \overline{c} . \rho_0'$, and by Corollary 2.20 $\vdash^{\partial s \ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \sigma_0'$. By the substitution lemma for (10) we have $\vdash_{\Uparrow}^{inst} \sigma_0' \le [\overline{g_1 \mapsto g_2}]\rho_2$ and by transitivity we have that $\vdash^{\partial s \ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le [\overline{g_1 \mapsto g_2}]\rho_2$. Now it cannot be that $\overline{g}_2 \in \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t$ because $\overline{g}_2 \notin ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get $\vdash^{\partial s \ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \forall \overline{g}_2 . [\overline{g_1 \mapsto g_2}]\rho_2$ or by dropping useless quantifiers and α-renaming $\vdash^{\partial s \ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \overline{\Gamma_2}(\rho_2)$ as required.

– Assume that $\rho_t = a$ and let $a \notin ftv(\Gamma_1)$. Then by equation (14) and Corollary 2.20 we get:

$$\vdash^{ol} \forall a . a \le \forall \overline{c} . \mathfrak{pr}(\sigma_0) \to \rho_0' \tag{26}$$

where

$$\overline{c} \notin ftv(\sigma_0, \Gamma_1, \Gamma_2) \tag{27}$$
$$\mathfrak{pr}(\sigma_0') = \forall \overline{c} . \rho_0' \tag{28}$$

By (26) and (27) and inversion on $\vdash^{ol}$ it must be that

$$\vdash^{ol} \tau_1 \to \tau_2 \le \mathfrak{pr}(\sigma_0) \to \rho_0' \tag{29}$$

Now yet one more inversion gives

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \le \tau_1 \tag{30}$$
$$\vdash^{ol} \tau_2 \le \rho_0' \tag{31}$$

From Corollary 2.20 and (30)

$$\vdash^{\partial s \ell} \sigma_0 \le \tau_1 \tag{32}$$

From (12) and the substitution lemma, we get

$$\Gamma_1 \vdash_{sd} t : \tau_1 \to \tau_2 \tag{33}$$

By the substitution lemma for (9) we have that $\Gamma_2 \vdash_{sd}^{poly} u : \overline{[g_1 \mapsto g_2]}\sigma_1$ and by induction hypothesis and (32) we have

$$\Gamma_1 \vdash_{sd}^{poly} u : \sigma_1' \vdash^{\partial s \ell} \sigma_1' \le \tau_1 \tag{34}$$

Then $\vdash^{inst} \tau_2 \le \tau_2$. We have all the premises of the rule APP and applying it gives us that $\Gamma_1 \vdash t \, u : \tau_2$. By (31) we get that

$$\vdash^{ol} \forall \overline{d} . \tau_2 \le \rho_0' \tag{35}$$

where $\overline{d} = ftv(\tau_2) - ftv(\Gamma_1)$. But now we know that $\overline{c} \notin ftv(\forall \overline{d} . \tau_2)$, because it must be that $ftv(\forall \overline{d}\,\overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t) \subseteq ftv(\Gamma_1)$, and by (27) $\overline{c} \notin ftv(\Gamma_1)$. Then we can apply rule SKOL to get that $\vdash^{ol} \forall \overline{d} . \tau_2 \le \forall \overline{c} . \rho_0'$, and by Corollary 2.20 $\vdash^{\partial s \ell} \forall \overline{d} . \tau_2 \le \sigma_0'$. By the substitution lemma for (10) we have $\vdash_{\Uparrow}^{inst} \sigma_0' \le \overline{[g_1 \mapsto g_2]}\rho_2$ and by transitivity we have that $\vdash^{\partial s \ell} \forall \overline{d} . \tau_2 \le \overline{[g_1 \mapsto g_2]}\rho_2$. Now it cannot be that $\overline{g}_2 \in \forall \overline{d} . \tau_2$ because $\overline{g}_2 \notin ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get $\vdash^{\partial s \ell} \forall \overline{d} . \tau_2 \le \forall \overline{g}_2 . \overline{[g_1 \mapsto g_2]}\rho_2$ or by dropping useless quantifiers and α-renaming $\vdash^{\partial s \ell} \forall \overline{d} . \tau_2 \le \overline{\Gamma_2}(\rho_2)$ as required.

• Case LET. In this case we have that $\Gamma_2 \vdash_{sd} \texttt{let } x = u \texttt{ in } t : \rho_2$, given that

$$\Gamma_2 \vdash_{sd}^{poly} u : \sigma \tag{36}$$
$$\Gamma_2, x : \sigma \vdash_{sd} t : \rho_2 \tag{37}$$

By induction hypothesis for (36) $\Gamma_1 \vdash_{sd} u : \sigma'$ such that $\vdash^{\partial s \ell} \sigma' \le \sigma$. By induction hypothesis for (37) we get $\Gamma_1, x : \sigma' \vdash_{sd} t : \rho_1$ such that $\vdash^{\partial s \ell} \overline{\Gamma_1, x : \sigma'}(\rho_1) \le \overline{\Gamma_2, x : \sigma}(\rho_2)$ or since σ' is generalised over Γ_1 and σ is generalised over Γ_2 this becomes $\vdash^{\partial s \ell} \overline{\Gamma_1}(\rho_1) \le \overline{\Gamma_2}(\rho_2)$ as required. Applying rule LET finishes the case.

- Case ANNOT. We have that $\Gamma_2 \vdash_{sd} (t::\sigma) : \rho_2$, given that

$$\Gamma_2 \vdash_{sd}^{poly} t : \sigma' \tag{38}$$

$$\vdash^{sub\sigma} \sigma' \leq \sigma \tag{39}$$

$$\vdash^{inst} \sigma \leq \rho_2 \tag{40}$$

From (38) and induction hypothesis $\Gamma_1 \vdash_{sd}^{poly} t : \sigma''$ such that $\vdash^{\partial s\ell} \sigma'' \leq \sigma'$. Then by (39) and transitivity of $\vdash^{\partial s\ell}$ we have $\vdash^{\partial s\ell} \sigma'' \leq \sigma$. Consider a renaming substitution $[\overline{b \mapsto d}]$, where $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2)$ and $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \sigma', \rho_2)$. Then by substitution lemma $\vdash^{inst} \sigma \leq [\overline{b \mapsto d}]\rho_2$. We need to show that $\vdash^{\partial s\ell} \overline{\Gamma_1}([\overline{b \mapsto d}]\rho_2) \leq \overline{\Gamma_2}([\overline{b \mapsto d}]\rho_2)$. But notice that $ftv([\overline{b \mapsto d}]\rho_2) - ftv(\Gamma_2) = \overline{d}$ and $\overline{d} \subseteq ftv([\overline{b \mapsto d}]\rho_2) - ftv(\Gamma_2))$ because $\overline{d}$ are fresh. The result then follows by rule SUB.

$\square$

Lemma 3.4. *Assume that $\vdash^{sub\sigma}$ is $\vdash^{\partial s\ell}$. If $\Gamma \vdash_{nsd} t : \sigma$ then $\Gamma \vdash_{sd} t : \rho$ and $\vdash^{\partial s\ell} \overline{\Gamma}(\rho) \leq \sigma$.*

Proof. By induction on height of the non syntax-directed derivations. We proceed by case analysis on the last rule used.

- Case INT. The result follows by rule INT and we know that $\vdash^{\partial s\ell} \texttt{Int} \leq \texttt{Int}$ by reflexivity.

- Case VAR. Int this case $\Gamma \vdash_{nsd} x : \sigma$, given that $x : \sigma \in \Gamma$. Then, assuming that $\sigma = \forall \overline{a}.\rho$ and without loss of generality $\overline{a} \notin ftv(\Gamma)$, we have that $\vdash^{inst} \sigma \leq \rho$ and by rule SUB we get that $\vdash^{\eta} \overline{\Gamma}(\rho) \leq \forall \overline{a}.\rho$ or by Corollary 2.31 $\vdash^{\partial s\ell} \overline{\Gamma}(\rho) \leq \forall \overline{a}.\rho$. By rule VAR we also get $\Gamma \vdash_{sd} x : \rho$ as required.

- Case ABS. We have that $\Gamma \vdash_{nsd} (\lambda x.t) : \tau \rightarrow \sigma$ given that $\Gamma, x : \tau \vdash_{nsd} t : \sigma$. By induction hypothesis there exists a ρ such that $\Gamma, x : \tau \vdash_{sd} t : \rho$ such that

$$\vdash^{\partial s\ell} \overline{\Gamma, x : \tau}(\rho) \leq \sigma \tag{1}$$

Let $\overline{b} = ftv(\rho) - ftv(\Gamma, \tau)$. By rule ABS we get that $\Gamma \vdash_{sd} (\lambda x.t) : \tau \rightarrow \rho$. We wish to show that $\vdash^{\partial s\ell} \overline{\Gamma}(\tau \rightarrow \rho) \leq \tau \rightarrow \sigma$. Let $\overline{a} = ftv(\tau \rightarrow \rho) - ftv(\Gamma)$. Split these variables in two sets $\overline{a} = \overline{a}_1 \overline{a}_2$, such that $\overline{a}_1 = \overline{a} - ftv(\tau)$ and $\overline{a}_2 = \overline{a} - \overline{a}_1$. Then it must be that $\overline{a}_1 = \overline{b}$. Then we have the following (using reflexivity and transitivity of $\vdash^{\partial s\ell}$):

$$
\begin{array}{rll}
\vdash^{\partial s\ell} \forall \overline{a}.\tau \rightarrow \rho & = & \forall \overline{a}_1 \overline{a}_2.\tau \rightarrow \rho \\
& \leq & \forall \overline{a}_2.(\forall \overline{a}_1.\tau) \rightarrow \forall \overline{a}_1.\rho \quad \text{(by DISTRIB and ALL)} \\
& \leq & (\forall \overline{a}_1.\tau) \rightarrow \forall \overline{a}_1.\rho \quad \text{(by SUB)} \\
& \leq & \tau \rightarrow \forall \overline{a}_1.\rho \quad \text{(by Lemma 2.25)} \\
& \leq & \tau \rightarrow \sigma \quad \text{(by (1) and FUN)}
\end{array}
$$

(The use of DISTRIB is essential for this derivation and it is a reason why the claim fails when $\vdash^{ol}$ is used instead of $\vdash^{\partial s\ell}$.)

- Case AABS. Similar argument as ABS taking into account that type annotations are closed.

- Case APP. In this case we have that $\Gamma \vdash_{nsd} t\, u : \sigma'$ given that

$$\Gamma \vdash_{nsd} t : (\sigma \rightarrow \sigma') \tag{2}$$

$$\Gamma \vdash_{nsd} u : \sigma \tag{3}$$

By induction hypothesis there exists ρ_t with

$$\Gamma \vdash_{sd} t : \rho_t \tag{4}$$

$$\vdash^{\partial s \ell} \overline{\Gamma}(\rho_t) \leq \sigma \to \sigma' \tag{5}$$

Moreover, by induction hypothesis for (3) we get

$$\Gamma \vdash_{sd} u : \rho_u \tag{6}$$

$$\vdash^{\partial s \ell} \overline{\Gamma}(\rho_u) \leq \sigma \tag{7}$$

In general there are two possible cases for ρ_t. It can either be a type variable $a \notin ftv(\Gamma)$ or it will be an arrow type.

– Assume that $\rho_t = \sigma_1^t \to \sigma_2^t$. Then by (5) $\vdash^{\partial s \ell} \forall \overline{a}.\sigma_1^t \to \sigma_2^t \leq \sigma \to \sigma'$, where $\overline{a} = ftv(\sigma_1^t, \sigma_2^t) - ftv(\Gamma)$. By the prenex corollary, Corollary 2.20

$$\vdash^{ol} \forall \overline{a}\,\overline{b}.\mathfrak{pr}(\sigma_1^t) \to \rho_2^t \leq \forall \overline{c}.\mathfrak{pr}(\sigma) \to \rho' \tag{8}$$

where the following are true:

$$\overline{b} \notin ftv(\sigma_1^t) \tag{9}$$

$$\overline{c} \notin ftv(\sigma, \Gamma, \sigma_1^t, \sigma_2^t) \tag{10}$$

$$\mathfrak{pr}(\sigma_2^t) = \forall \overline{b}.\rho_2^t \tag{11}$$

$$\mathfrak{pr}(\sigma') = \forall \overline{c}.\rho' \tag{12}$$

By (8) and inversion on $\vdash^{ol}$ we get

$$
\begin{aligned}
& \vdash^{ol} \forall \overline{a}\,\overline{b}.\mathfrak{pr}(\sigma_1^t) \to \rho_2^t \leq \mathfrak{pr}(\sigma) \to \rho' && \text{(by inversion)} \\
\Rightarrow \quad & \vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}](\mathfrak{pr}(\sigma_1^t) \to \rho_2^t) \leq \mathfrak{pr}(\sigma) \to \rho' && \text{(by inversion)} \\
\Rightarrow \quad & \vdash^{ol} [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \to [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \leq \mathfrak{pr}(\sigma) \to \rho'
\end{aligned}
$$

From the last equation, by inversion we get that

$$\vdash^{ol} \mathfrak{pr}(\sigma) \leq [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \tag{13}$$

$$\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \leq \rho' \tag{14}$$

From Corollary (2.20) and (13)

$$\vdash^{ol} \sigma \leq [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{15}$$

Now, from (4) and the substitution lemma, we get

$$\Gamma \vdash_{sd} t : [\overline{a \mapsto \tau_a}](\sigma_1^t \to \sigma_2^t) \tag{16}$$

By transitivity of $\vdash^{\partial s \ell}$, (15) and (7) we have

$$\vdash^{\partial s \ell} \overline{\Gamma}(\rho_u) \leq [\overline{a \mapsto \tau_a}]\sigma_2^t \tag{17}$$

By (15), (6), and (17) we can apply rule APP to get that $\Gamma \vdash_{sd} t\ u : \rho_r$ such that $\vdash^{inst} [\overline{a \mapsto \tau_a}]\sigma_2^t \leq \rho_r$ and just pick ρ_r such that all the quantifiers of $[\overline{a \mapsto \tau_a}]\sigma_2^t$ are replaced by variables not in Γ. Then it will be the case that $\overline{\Gamma}(\rho_r) = \overline{\Gamma}([\overline{a \mapsto \tau_a}]\sigma_2^t)$. Consequently to finish the case we need to show that

$$\vdash^{\partial s \ell} \forall \overline{d}.[\overline{a \mapsto \tau_a}]\sigma_2^t \leq \sigma' \tag{18}$$

where $\overline{d} = ftv(\rho_2^t, \overline{\tau_a}) - ftv(\Gamma)$. By (14) we have

$$\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \leq \rho'$$
$$\Rightarrow \quad \vdash^{ol} \forall \overline{b}.[\overline{a \mapsto \tau_a}]\rho_2^t \leq \rho'$$
$$\Rightarrow \quad \vdash^{ol} \forall \overline{db}.[\overline{a \mapsto \tau_a}]\rho_2^t \leq \rho'$$

But now we know that $\overline{c} \notin ftv(\forall \overline{db}.[\overline{a \mapsto \tau_a}]\rho_2^t)$, because it must be that $ftv(\forall \overline{db}.[\overline{a \mapsto \tau_a}]\rho_2^t) \subseteq$ $ftv(\Gamma)$, and by (10) $\overline{c} \notin ftv(\Gamma)$. Then we can apply rule SKOL to get that $\vdash^{ol} \forall \overline{db}.[\overline{a \mapsto \tau_a}]\rho_2^t \leq$ $\forall \overline{c}.\rho'$, and by Corollary 2.20 $\vdash^{\partial s\ell} \forall \overline{d}.[\overline{a \mapsto \tau_a}]\sigma_2^t \leq \sigma'$ as required.

- Assume that $\rho_t = a$. The argument is essentially the same. We give it more briefly. Here it must be that $a \notin ftv(\Gamma)$ so that $\overline{\Gamma}(a) = \forall a.a$; otherwise (5) cannot be derivable. Then, by (5) we have that $\vdash^{\partial s\ell} \forall a.a \leq \sigma \to \sigma'$, which, by Corollary 2.20 gives $\vdash^{ol} \forall a.a \leq \forall \overline{c}.\mathfrak{pr}(\sigma) \to \rho'$ where $\mathfrak{pr}(\sigma') = \forall \overline{c}.\rho'$ and assume that $\overline{c} \notin ftv(\sigma, \Gamma)$. By inversion $\vdash^{ol} \forall a.a \leq \mathfrak{pr}(\sigma) \to \rho'$, and there exist τ_1, τ_2 with $\vdash^{ol} [a \mapsto (\tau_1 \to \tau_2)]a \leq \mathfrak{pr}(\sigma) \to \rho'$ or by one more inversion

$$\vdash^{ol} \mathfrak{pr}(\sigma) \leq \tau_1 \tag{19}$$

$$\vdash^{ol} \tau_2 \leq \rho' \tag{20}$$

By the substitution lemma, Lemma 3.2, $\Gamma \vdash_{sd} t : \tau_1 \to \tau_2$. Moreover by Corollary 2.20 and (19) we get that $\vdash^{\partial s\ell} \sigma \leq \tau_1$. Therefore by (7) and transitivity of $\vdash^{\partial s\ell}$, $\vdash^{\partial s\ell} \overline{\Gamma}(\rho_u) \leq \tau_1$. Finally we can apply rule APP to get the result. We need to show that $\vdash^{\partial s\ell} \overline{\Gamma}(\tau_2) \leq \sigma'$. But we have that $\vdash^{ol} \overline{\Gamma}(\tau_2) \leq \rho'$ and since we assumed that $\overline{c} \notin ftv(\Gamma)$, $\overline{c} \notin ftv(\overline{\Gamma}(\tau_2))$, therefore we can apply rule SKOL to get $\vdash^{ol} \overline{\Gamma}(\tau_2) \leq \forall \overline{c}.\rho'$ and by this and Corollary 2.20 we are done.

- Case LET. We have that $\Gamma \vdash_{nsd} \texttt{let } x = u \texttt{ in } t : \sigma'$, given that

$$\Gamma \vdash_{nsd} u : \sigma \tag{21}$$

$$\Gamma, x : \sigma \vdash_{nsd} t : \sigma' \tag{22}$$

By induction hypothesis $\Gamma \vdash_{sd} u : \rho$ such that $\vdash^{\partial s\ell} \overline{\Gamma}(\rho) \leq \sigma$. Again by induction hypothesis $\Gamma, x : \sigma \vdash_{sd} t : \rho'$ such that $\vdash^{\partial s\ell} \overline{\Gamma, x : \sigma}(\rho') \leq \sigma'$. By Lemma 3.3 we get that $\Gamma, x : \overline{\Gamma}(\rho) \vdash_{sd} t : \rho''$ such that $\vdash^{\partial s\ell} \overline{\Gamma, x : \overline{\Gamma}(\rho)}(\rho'') \leq \overline{\Gamma, x : \sigma}(\rho')$. But $\overline{\Gamma, x : \overline{\Gamma}(\rho)}(\rho'') = \overline{\Gamma}(\rho'')$ and $\vdash^{\partial s\ell} \overline{\Gamma, x : \sigma}(\rho') \leq \sigma'$ by transitivity of $\vdash^{\partial s\ell}$. The result follows by application of rule LET.

- Case ANNOT. In this case we have that $\Gamma \vdash_{nsd} (t::\sigma) : \sigma$, given that $\Gamma \vdash_{nsd} t : \sigma$. By induction hypothesis there exists a ρ such that $\Gamma \vdash_{sd} t : \rho$ and $\vdash^{\partial s\ell} \overline{\Gamma}(\rho) \leq \sigma$. Then is the case that $\Gamma \vdash_{sd}^{poly} t : \overline{\Gamma}(\rho)$. Moreover if $\sigma = \forall \overline{a}.\rho_1$ and without loss of generality $\overline{a} \notin ftv(\Gamma)$, then $\vdash^{inst} \sigma \leq \rho_1$ and we are done by applying rule ANNOT. Moreover it is easy to confirm that $\vdash^{\partial s\ell} \overline{\Gamma}(\rho_1) \leq \forall \overline{a}.\rho_1$, since we assumed $\overline{a} \notin ftv(\Gamma)$.

$\square$

The above claim fails when $\vdash^{sub\sigma}$ is $\vdash^{ol}$. For example it is derivable in the non-syntax directed system that

$$\vdash_{nsd} \texttt{let } f = (\backslash x.\backslash y.y) \texttt{ in } f : \texttt{Int} \to \forall b.b \to b$$

but in the syntax-directed system we can only get that

$$\vdash_{sd}^{poly} \texttt{let } f = (\backslash x.\backslash y.y) \texttt{ in } f : \forall ab.a \to b \to b$$

and it is not the case that $\vdash^{ol} \forall ab.a \to b \to b \leq \texttt{Int} \to \forall b.b \to b$, even though it is the case that

$$\vdash^{\partial s\ell} \forall ab.a \to b \to b \leq \texttt{Int} \to \forall b.b \to b$$

$$\text{Rho-types} \quad \rho \ ::= \ \tau \mid \sigma \to \sigma$$

$$\boxed{\Gamma \vdash_\delta t : \rho} \qquad \delta ::= \ \Uparrow \mid \Downarrow$$

$$\frac{}{\Gamma \vdash_\delta i : \texttt{Int}} \ \text{INT} \qquad \frac{\vdash^{inst}_\delta \sigma \le \rho}{\Gamma, (x : \sigma) \vdash_\delta x : \rho} \ \text{VAR}$$

$$\frac{\Gamma, (x : \tau) \vdash_\Uparrow t : \rho}{\Gamma \vdash_\Uparrow (\backslash x . t) : (\tau \to \rho)} \ \text{ABS1} \qquad \frac{\Gamma, (x : \sigma_a) \vdash^{poly}_\Downarrow t : \sigma_r}{\Gamma \vdash_\Downarrow (\backslash x . t) : (\sigma_a \to \sigma_r)} \ \text{ABS2}$$

$$\frac{\Gamma, (x : \sigma) \vdash_\Uparrow t : \rho}{\Gamma \vdash_\Uparrow (\backslash (x :: \sigma) . t) : (\sigma \to \rho)} \ \text{AABS1} \qquad \frac{\vdash^{sub\sigma} \sigma_a \le \sigma_x \quad \Gamma, (x : \sigma_x) \vdash^{poly}_\Downarrow t : \sigma_r}{\Gamma \vdash_\Downarrow (\backslash (x :: \sigma_x) . t) : (\sigma_a \to \sigma_r)} \ \text{AABS2}$$

$$\frac{\Gamma \vdash_\Uparrow t : (\sigma \to \sigma') \quad \Gamma \vdash^{poly}_\Downarrow u : \sigma \quad \vdash^{inst}_\delta \sigma' \le \rho}{\Gamma \vdash_\delta t \ u : \rho} \ \text{APP}$$

$$\frac{\begin{array}{c} \Gamma \vdash^{poly}_\Downarrow t : \sigma \\ \vdash^{inst}_\delta \sigma \le \rho \end{array}}{\Gamma \vdash_\delta (t :: \sigma) : \rho} \ \text{ANNOT} \qquad \frac{\begin{array}{c} \Gamma \vdash^{poly}_\Uparrow u : \sigma \\ \Gamma, x : \sigma \vdash_\delta t : \rho \end{array}}{\Gamma \vdash_\delta \texttt{let } x = u \texttt{ in } t : \rho} \ \text{LET}$$

$$\boxed{\Gamma \vdash^{poly}_\delta t : \sigma}$$

$$\frac{\begin{array}{c} \overline{a} = \mathit{ftv}(\rho) - \mathit{ftv}(\Gamma) \\ \Gamma \vdash_\Uparrow t : \rho \end{array}}{\Gamma \vdash^{poly}_\Uparrow t : \forall \overline{a} . \rho} \ \text{GEN1} \qquad \frac{\begin{array}{c} \overline{a} \notin \mathit{ftv}(\Gamma) \\ \Gamma \vdash_\Downarrow t : \rho \end{array}}{\Gamma \vdash^{poly}_\Downarrow t : \forall \overline{a} . \rho} \ \text{GEN2}$$

$$\boxed{\vdash^{inst}_\delta \sigma \le \rho}$$

$$\frac{}{\vdash^{inst}_\Uparrow \forall \overline{a} . \rho \ \le \ [\overline{a} \mapsto \overline{\tau}] \, \rho} \ \text{INST1} \qquad \frac{\vdash^{sub\sigma} \sigma \ \le \ \rho}{\vdash^{inst}_\Downarrow \sigma \ \le \ \rho} \ \text{INST2}$$

Figure 9: Bidirectional higher-rank type system

3.2 Bidirectional type system (first version)

In this section we give properties of the type system of Figure 9. Notice that this is not the final version of the type system as it performs only shallow skolemisation in the rule GEN2. Nevertheless it is worth studying its properties. We later extend it to the final version which also appears in the main paper.

Lemma 3.5. *If* $\vdash^{sub\sigma} \sigma \le \rho$ *then* $\vdash^{inst}_\Downarrow \sigma \le \rho$.

Proof. Directly follows by rule INST2. □

Lemma 3.6. *If* $\vdash^{inst}_{\Uparrow} \sigma \le \rho$ *then* $\vdash^{inst}_{\Downarrow} \sigma \le \rho$.

Proof. By inversion it must be that $\sigma = \forall \overline{a}.\rho_1$ and $\rho = [\overline{a \mapsto \tau}]\rho_1$. We need to show that $\vdash^{inst}_{\Downarrow} \forall \overline{a}.\rho_1 \le [\overline{a \mapsto \tau}]\rho_1$, or by rule INST2 $\vdash^{sub\sigma} \forall \overline{a}.\rho_1 \le [\overline{a \mapsto \tau}]\rho_1$. By rule SUB $\vdash^{\eta} \forall \overline{a}.\rho_1 \le [\overline{a \mapsto \tau}]\rho_1$ and by Corollary 2.31 $\vdash^{\partial s\ell} \forall \overline{a}.\rho_1 \le [\overline{a \mapsto \tau}]\rho_1$. On the other hand, if $\vdash^{sub\sigma}$ is $\vdash^{ol}$ the result follows by an application of SPEC and reflexivity, Lemma 2.2. □

Lemma 3.7. *If* $\vdash^{ol} \sigma_1 \le \sigma_2$ *and* $\vdash^{inst}_{\Uparrow} \sigma_2 \le \rho_2$ *then* $\exists \rho_1. \vdash^{inst}_{\Uparrow} \sigma_1 \le \rho_1$ *and* $\vdash^{ol} \rho_1 \le \rho_2$.

Proof. Assume that $\sigma_1 = \forall \overline{a}.\rho_{11}$ and $\sigma_2 = \forall \overline{b}.\rho_{22}$. Without loss of generality assume that $\overline{a} \notin ftv(\sigma_2)$ and $\overline{b} \notin ftv(\sigma_1)$. Then it must be that

$$\vdash^{ol} \forall \overline{a}.\rho_{11} \le \forall \overline{b}.\rho_{22} \tag{1}$$

By inversion on the assumption we get that

$$\rho_2 = [\overline{b \mapsto \tau_b}]\rho_{22} \tag{2}$$

By (1) and inversion it must be that $\vdash^{ol} \forall \overline{a}.\rho_{11} \le \rho_{22}$. Then, by the substitution lemma $\vdash^{ol} \forall \overline{a}.\rho_{11} \le [\overline{b \mapsto \tau_b}]\rho_{22}$. Then by inversion again[4] $\vdash^{ol} [\overline{a \mapsto \tau_a}]\rho_{11} \le [\overline{b \mapsto \tau_b}]\rho_{22}$ for some $\overline{\tau_a}$. Taking $\rho_1 = [\overline{a \mapsto \tau_a}]\rho_{11}$ finishes the proof. □

Lemma 3.8. *If* $\vdash^{ol} \sigma_1 \le \sigma_2$ *and* $\vdash^{inst}_{\Downarrow} \sigma_2 \le \rho_2$ *then* $\vdash^{inst}_{\Downarrow} \sigma_1 \le \rho_2$.

Proof. Follows by inversion of $\vdash^{inst}$ and transitivity of $\vdash^{sub\sigma}$. □

Lemma 3.9. *Independently of whether* $\vdash^{sub\sigma}$ *is* $\vdash^{ol}$ *or* $\vdash^{\partial s\ell}$, *if* $\vdash^{ol} \Gamma_1 \le \Gamma_2$ *pointwise then the following are true:*

1. *If* $\Gamma_2 \vdash_{\Uparrow} t : \rho_2$ *then* $\exists \rho_1. \Gamma_1 \vdash_{\Uparrow} t : \rho_1$ *and* $\vdash^{ol} \rho_1 \le \rho_2$.

2. *If* $\Gamma_2 \vdash_{\Downarrow} t : \rho_1$ *and* $\vdash^{ol} \rho_1 \le \rho_2$ *then* $\Gamma_1 \vdash_{\Downarrow} t : \rho_2$.

3. *If* $\Gamma_2 \vdash^{poly}_{\Uparrow} t : \sigma_2$ *then* $\exists \sigma_1. \Gamma_1 \vdash^{poly}_{\Uparrow} t : \sigma_1$ *and* $\vdash^{ol} \sigma_1 \le \sigma_2$.

4. *If* $\Gamma_2 \vdash^{poly}_{\Downarrow} t : \sigma_1$ *and* $\vdash^{ol} \sigma_1 \le \sigma_2$ *then* $\Gamma_1 \vdash^{poly}_{\Downarrow} t : \sigma_2$.

Proof. We prove the four claims simultaneously by induction on the height of the typing derivations. For each claim we asssume that all are true for derivations of smaller height.

For the first part we have the following cases.

- Case INT. Follows by the same rule.

- Case VAR. In this case $\Gamma_2 \vdash_{\Uparrow} x : \rho_2$ given that $x : \sigma_2 \in \Gamma_2$ and $\vdash^{inst}_{\Uparrow} \sigma_2 \le \rho_2$. It must be that $x : \sigma_1 \in \Gamma_1$, such that $\vdash^{ol} \sigma_1 \le \sigma_2$. By Lemma 3.7 there exists ρ_1 such that $\vdash^{inst}_{\Uparrow} \sigma_1 \le \rho_1$ with $\vdash^{ol} \rho_1 \le \rho_2$. and by applying rule VAR we are done.

---

[4]Notice that this inversion would fail if we were using $\vdash^{\partial s\ell}$ instead.

- Case ABS1. Here $\Gamma_2 \vdash_\Uparrow (\backslash x.t) : \tau \to \rho$, given that $\Gamma_2, x : \tau \vdash_\Uparrow t : \rho$. By induction hypothesis $\Gamma_1, x : \tau \vdash_\Downarrow t : \rho'$ such that $\vdash^{ol} \rho' \le \rho$. By rule ABS1 we get $\Gamma_1 \vdash_\Uparrow (\backslash x.t) : \tau \to \rho'$. And by Lemma 2.2 and rule FUN $\vdash^{ol} \tau \to \rho' \le \tau \to \rho$.

- Case AABS1. In this case $\Gamma_2 \vdash_\Uparrow \backslash(x::\sigma).t : \sigma \to \rho$, given that $\Gamma_2, x : \sigma \vdash_\Uparrow t : \rho$. By induction hypothesis $\Gamma_1, x : \sigma \vdash_\Uparrow t : \rho'$ for some ρ' with $\vdash^{ol} \rho' \le \rho$ and by rule AABS1 we get that $\Gamma_1 \vdash_\Uparrow \backslash(x::\sigma).t : \sigma \to \rho'$ and by rule FUN $\vdash^{ol} \sigma \to \rho' \le \sigma \to \rho$.

- Case APP. In this case

$$\Gamma_2 \vdash_\Uparrow t\ u : \rho_2 \tag{1}$$

$$\Gamma_2 \vdash_\Uparrow t : (\sigma \to \sigma') \tag{2}$$

$$\Gamma_2 \vdash^{poly}_\Downarrow u : \sigma \tag{3}$$

$$\vdash^{inst}_\Uparrow \sigma' \le \rho_2 \tag{4}$$

  By induction hypothesis there exists ρ' such that $\Gamma_1 \vdash_\Uparrow t : \rho'$ and $\vdash^{ol} \sigma \to \sigma' \le \rho'$. By inversion it must be that $\rho' = \sigma_1 \to \sigma'_1$ such that $\vdash^{ol} \sigma \le \sigma_1$ and $\vdash^{ol} \sigma'_1 \le \sigma'$. From this and (3) and induction hypothesis it must be that $\Gamma_1 \vdash^{poly}_\Downarrow u : \sigma_1$ and by Lemma 3.7 there exists a ρ_1 with $\vdash^{ol} \rho_1 \le \rho_2$ such that $\vdash^{inst}_\Uparrow \sigma'_1 \le \rho_1$. By applying rule APP we are done.

- Case ANNOT. Here $\Gamma_2 \vdash_\Uparrow (t::\sigma) : \rho_2$, given that $\Gamma_2 \vdash^{poly}_\Downarrow t : \sigma$ and $\vdash^{inst}_\Uparrow \sigma \le \rho_2$. By reflexivity and induction hypothesis $\Gamma_1 \vdash^{poly}_\Downarrow t : \sigma$. Applying rule ANNOT again gives the result, since $\vdash^{ol} \rho_2 \le \rho_2$.

- Case LET. Finally $\Gamma_2 \vdash_\Uparrow \texttt{let } x \texttt{ = } u \texttt{ in } t : \rho_2$, given that $\Gamma_2 \vdash^{poly}_\Uparrow u : \sigma$ and $\Gamma_2, x : \sigma \vdash_\Uparrow t : \rho_2$. By induction hypothesis $\Gamma_1 \vdash^{poly}_\Uparrow u : \sigma'$ such that $\vdash^{ol} \sigma' \le \sigma$. Then by induction hypothesis again $\Gamma_1, x : \sigma' \vdash_\Uparrow t : \rho_1$ for some ρ_1 with $\vdash^{ol} \rho_1 \le \rho_2$. Applying LET finishes the case.

For the second part we have the following cases.

- Case INT. By inversion it must be also that $\rho_2 = \texttt{Int}$. Then the case follows by INT.

- Case VAR. In this case $\Gamma_2 \vdash_\Downarrow x : \rho_1$ given that $x : \sigma_2 \in \Gamma_2$ and $\vdash^{inst}_\Downarrow \sigma_2 \le \rho_1$. It must be that $x : \sigma_1 \in \Gamma_1$, such that $\vdash^{ol} \sigma_1 \le \sigma_2$. By Lemma 3.8 $\vdash^{inst}_\Downarrow \sigma_1 \le \rho_1$ and by transitivity of $\vdash^{sub\sigma}$ $\vdash^{inst}_\Downarrow \sigma_1 \le \rho_2$. The result follows by rule VAR.

- Case ABS2. Here $\Gamma_2 \vdash_\Downarrow (\backslash x.t) : \sigma_a \to \sigma_r$, given that $\Gamma_2, x : \sigma_a \vdash^{poly}_\Downarrow t : \sigma_r$. We have that $\vdash^{ol} \sigma_a \to \sigma_r \le \rho_2$ for some ρ_2. By inversion $\rho_2 = \sigma^a_2 \to \sigma^r_2$ with

$$\vdash^{ol} \sigma^a_2 \le \sigma_a \tag{5}$$

$$\vdash^{ol} \sigma_r \le \sigma^r_2 \tag{6}$$

  Then by induction hypothesis $\Gamma_1, x : \sigma^a_2 \vdash^{poly}_\Downarrow t : \sigma^r_2$. The result follows by applying rule ABS2 again.

- Case AABS2. In this case $\Gamma_2 \vdash_\Downarrow \backslash(x::\sigma_x).t : \sigma_a \to \sigma_r$, given that $\Gamma_2, x : \sigma_x \vdash^{poly}_\Downarrow t : \sigma_r$ and $\vdash^{sub\sigma} \sigma_a \le \sigma_x$. We have that $\vdash^{ol} \sigma_a \to \sigma_r \le \rho_2$ for some ρ_2. By inversion $\rho_2 = \sigma^a_2 \to \sigma^r_2$ with

$$\vdash^{ol} \sigma^a_2 \le \sigma_a \tag{7}$$

$$\vdash^{ol} \sigma_r \le \sigma^r_2 \tag{8}$$

  By transitivity of $\vdash^{sub\sigma}$ and the fact that $\vdash^{ol}$ is a subset of $\vdash^{sub\sigma}$, and (7) we get that $\vdash^{sub\sigma} \sigma^a_2 \le \sigma_x$. By induction hypothesis $\Gamma_1, x : \sigma_x \vdash^{poly}_\Downarrow t : \sigma^r_2$ and by rule AABS2 we get the result.

- Case APP. In this case

$$\Gamma_2 \vdash_\Downarrow t\ u : \rho_1 \tag{9}$$

$$\Gamma_2 \vdash_\Uparrow t : (\sigma \to \sigma') \tag{10}$$

$$\Gamma_2 \vdash^{poly}_\Downarrow u : \sigma \tag{11}$$

$$\vdash^{inst}_\Downarrow \sigma' \le \rho_1 \tag{12}$$

By induction hypothesis there exists ρ' such that $\Gamma_1 \vdash_\Uparrow t : \rho'$ and $\vdash^{ol} \rho' \le \sigma \to \sigma'$. By inversion it must be that $\rho' = \sigma_1 \to \sigma'_1$ such that $\vdash^{ol} \sigma \le \sigma_1$ and $\vdash^{ol} \sigma'_1 \le \sigma'$. From this and (11) and induction hypothesis it must be that $\Gamma_1 \vdash^{poly}_\Downarrow u : \sigma_1$ and by transitivity of $\vdash^{sub\sigma}$, $\vdash^{inst}_\Uparrow \sigma'_1 \le \rho_2$. By applying rule APP we are done.

- Case ANNOT. Here $\Gamma_2 \vdash_\Downarrow (t::\sigma) : \rho_1$, given that $\Gamma_2 \vdash^{poly}_\Downarrow t : \sigma$ and $\vdash^{inst}_\Downarrow \sigma \le \rho_1$. By reflexivity and induction hypothesis $\Gamma_1 \vdash^{poly}_\Downarrow t : \sigma$. By transitivity of $\vdash^{sub\sigma}$ we have that $\vdash^{inst}_\Downarrow \sigma \le \rho_2$, and by rule ANNOT we are done.

- Case LET. Here $\Gamma_2 \vdash_\Downarrow \texttt{let } x = u \texttt{ in } t : \rho_1$, given that $\Gamma_2 \vdash^{poly}_\Uparrow u : \sigma$ and $\Gamma_2, x : \sigma \vdash_\Downarrow t : \rho_1$. By induction hypothesis $\Gamma_1 \vdash^{poly}_\Uparrow u : \sigma'$ so that $\vdash^{ol} \sigma' \le \sigma$. Then by induction hypothesis again $\Gamma_1, x : \sigma' \vdash_\Downarrow t : \rho_2$ and applying LET finishes the case.

For the third part we have by inversion that $\sigma_2 = \forall \overline{b}.\rho_2$, such that $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2)$ and

$$\Gamma_2 \vdash_\Uparrow t : \rho_2 \tag{13}$$

Instead of using the induction hypothesis directly[5] consider a renaming substitution $[\overline{b \mapsto d}]$ such that $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \rho_2)$. Then by the substitution lemma, Lemma 3.13 on (13 we get $\Gamma_2 \vdash_\Uparrow t : [\overline{b \mapsto d}]\rho_2$ with the same height. Then we can apply the induction hypothesis to get that $\Gamma_1 \vdash_\Uparrow t : \rho_1$, such that $\vdash^{ol} \rho_1 \le [\overline{b \mapsto d}]\rho_2$. By rule SPEC

$$\vdash^{ol} \forall \overline{a}.\rho_1 \le [\overline{b \mapsto d}]\rho_2 \tag{14}$$

where $\overline{a} = ftv(\rho_1) - ftv(\Gamma_1)$. Now we claim that $\overline{d} \notin ftv(\forall \overline{a}.\rho_1)$. Suppose by contradiction that exists a $d \in \overline{d}$ such that $d \in ftv(\forall \overline{a}.\rho_1) = ftv(\rho_1) \cap ftv(\Gamma_1)$. Then it must be that $d \in ftv(\Gamma_1)$, a contradiction. By rule SKOL then and (14) we get $\vdash^{ol} \forall \overline{a}.\rho_1 \le \forall \overline{d}.[\overline{b \mapsto d}]\rho_2$, or equivalently $\vdash^{ol} \forall \overline{a}.\rho_1 \le \forall \overline{b}.\rho_2$ as required.

For the fourth part assume that $\sigma_1 = \forall \overline{a}.\rho_1$ and $\sigma_2 = \forall \overline{b}.\rho_2$. Without loss of generality assume that $\overline{b} \notin ftv(\sigma_1, \Gamma_1)$. Then we have that $\vdash^{ol} \forall \overline{a}.\rho_1 \le \forall \overline{b}.\rho_2$ and by inversion $\vdash^{ol} \forall \overline{a}.\rho_1 \le \rho_2$. By inversion again[6] we get

$$\vdash^{ol} [\overline{a \mapsto \tau}]\rho_1 \le \rho_2 \tag{15}$$

for some $\overline{\tau}$. We know that $\Gamma_2 \vdash_\Downarrow \rho_1$ and $\overline{a} \notin ftv(\Gamma_2)$. Then by the substitution lemma, Lemma 3.13 $\Gamma_2 \vdash_\Downarrow t : [\overline{a \mapsto \tau}]\rho_1$. From (15) and by induction hypothesis $\Gamma_1 \vdash_\Downarrow t : \rho_2$. By rule GEN2 we get the result.

$\square$

Notice that the property holds independently of which relation the type system actually uses. However it fails when the two types are related in $\vdash^{dsk}$ instead of $\vdash^{ol}$.

False Claim 3.10. *If* $\Gamma \vdash_\Downarrow t : \rho_1$ *and* $\vdash^{dsk} \rho_1 \le \rho_2$ *then* $\Gamma \vdash_\Downarrow t : \rho_2$.

---

[5] Induction hypothesis would give that $\Gamma_1 \vdash_\Uparrow : \rho_1$ such that $\vdash^{ol} \rho_1 \le \rho_2$, but in general it is not true that if $\vdash^{ol} \Gamma_1 \le \Gamma_2$ and $\vdash^{ol} \rho_1 \le \rho_2$ then $\vdash^{ol} \overline{\Gamma_1}(\rho_1) \le \overline{\Gamma_2}(\rho_2)$. As a counterexample consider $\Gamma_1 = (x : (a \to a) \to \texttt{Int})$, $\Gamma_2 = (x : (\forall a.\,a \to a) \to \texttt{Int})$ and $\rho_1 = \rho_2 = a \to a$.

[6] Note that this step would fail if we were in $\vdash^{dsk}$.

Proof. Here's a counterexample. Consider $\Gamma = u : \text{Int}$, $\sigma_1 = \forall a.\, a \rightarrow \forall b.\, b \rightarrow \forall c.\, b \rightarrow c$, $\sigma_2 = \text{Int} \rightarrow \forall c.\, \text{Int} \rightarrow c$, and $\sigma_3 = \forall abc.\, a \rightarrow b \rightarrow b \rightarrow c$. Then it is derivable that $\Gamma \vdash_{\Downarrow} (\backslash x.\, x\ u) : (\sigma_1 \rightarrow \sigma_2)$ but it is not derivable that $\Gamma \vdash_{\Downarrow} (\backslash x.\, x\ u) : (\sigma_3 \rightarrow \sigma_2)$, although $\vdash^{\partial s\ell} \sigma_1 \rightarrow \sigma_2 \leq \sigma_3 \rightarrow \sigma_2$. Notice that this property fails again independently of $\vdash^{sub\sigma}$. □

Lemma 3.11.

1. *If* $\Gamma \vdash_{\Uparrow} t : \rho$ *then* $\Gamma \vdash_{\Downarrow} t : \rho$.

2. *If* $\Gamma \vdash^{poly}_{\Uparrow} t : \sigma$ *then* $\Gamma \vdash^{poly}_{\Downarrow} t : \sigma$.

Proof. We prove the two claims simultaneously by induction on the height of the derivations. We proceed with case analysis on the last rule used in each derivation. For the first part we have the following cases to consider.

- Case INT. Follows by the same rule.

- Case VAR. In this case $\Gamma \vdash_{\Uparrow} x : \rho$ given that $x : \sigma \in \Gamma$ and $\vdash^{inst}_{\Uparrow} \sigma \leq \rho$. By Lemma 3.6 $\vdash^{inst}_{\Downarrow} \sigma \leq \rho$ and by applying rule VAR we are done.

- Case ABS1. Here $\Gamma \vdash_{\Uparrow} (\backslash x.\, t) : \tau \rightarrow \rho$, given that $\Gamma, x : \tau \vdash_{\Uparrow} t : \rho$. By induction hypothesis $\Gamma, x : \tau \vdash_{\Downarrow} t : \rho$ and by rule GEN2 $\Gamma, x : \tau \vdash^{poly}_{\Downarrow} t : \rho$. From this and rule ABS2 $\Gamma \vdash_{\Downarrow} (\backslash x.\, t) : \tau \rightarrow \rho$.

- Case AABS1. In this case $\Gamma \vdash_{\Uparrow} \backslash(x::\sigma).\, t : \sigma \rightarrow \rho$, given that $\Gamma, x : \sigma \vdash_{\Uparrow} t : \rho$. By induction hypothesis $\Gamma, x : \sigma \vdash_{\Downarrow} t : \rho$ and by rule GEN2 $\Gamma, x : \sigma \vdash^{poly}_{\Downarrow} t : \rho$. Moreover by reflexivity $\vdash^{sub\sigma} \sigma \leq \sigma$ and by applying rule AABS2 we are done.

- Case APP. In tis case $\Gamma \vdash_{\Uparrow} t\ u : \rho$, given that $\Gamma \vdash_{\Uparrow} t : (\sigma \rightarrow \sigma')$, $\Gamma \vdash^{poly}_{\Downarrow} u : \sigma$, and $\vdash^{inst}_{\Uparrow} \sigma' \leq \rho$. By Lemma 3.6 $\vdash^{inst}_{\Downarrow} \sigma' \leq \rho$ and we get the result by applying rule APP.

- Case ANNOT. Here $\Gamma \vdash_{\Uparrow} (t::\sigma) : \rho$, given that $\Gamma \vdash^{poly}_{\Downarrow} t : \sigma$ and $\vdash^{inst}_{\Uparrow} \sigma \leq \rho$. By Lemma 3.6 $\vdash^{inst}_{\Downarrow} \sigma \leq \rho$ and we get the result by rule ANNOT.

- Case LET. Finally $\Gamma \vdash_{\Uparrow} \text{let } x = u \text{ in } t : \rho$, given that $\Gamma \vdash^{poly}_{\Uparrow} u : \sigma$ and $\Gamma, x : \sigma \vdash_{\Uparrow} t : \rho$. By induction hypothesis $\Gamma, x : \sigma \vdash_{\Downarrow} t : \rho$ and the result follows by rule LET.

The second part can be derived with rule GEN1. It is $\Gamma \vdash^{poly}_{\Uparrow} t : \forall \overline{a}.\, \rho$, with $\overline{a} = ftv(\rho) - ftv(\Gamma)$ and $\Gamma \vdash_{\Uparrow} t : \rho$. By induction hypothesis $\Gamma \vdash_{\Downarrow} t : \rho$ and by using rule GEN2 we get $\Gamma \vdash^{poly}_{\Downarrow} t : \forall \overline{a}.\, \rho$ as required. □

Lemma 3.12 (Instantiation Substitution). *If* $\vdash^{inst}_{\delta} \sigma \leq \rho$ *then* $\vdash^{inst}_{\delta} S\sigma \leq S\rho$.

Proof. For $\delta = \Downarrow$ the result follows directly by the substitution property for $\vdash^{sub\sigma}$. For $\delta = \Uparrow$ by inversion $\sigma = \forall \overline{a}.\, \rho_1$ and $\rho = [\overline{a \mapsto \tau}]\rho_1$. Assume without loss of generality that $\overline{a} \notin vars(S)$. Then $S\sigma = \forall \overline{a}.\, S\rho_1$ and $S\rho = [\overline{a \mapsto S\tau}]S\rho_1$. The result follows by rule INST1. □

Lemma 3.13 (Substitution).

1. *If* $\Gamma \vdash_{\delta} t : \rho$ *then* $S\Gamma \vdash_{\delta} t : S\rho$.

2. *If* $\Gamma \vdash^{poly}_{\delta} t : \sigma$ *then* $S\Gamma \vdash^{poly}_{\delta} t : S\sigma$.

Proof. We prove all claims simultaneously by induction on the height of the derivations. For each of the claims we assume that all claims hold for derivations of smaller height.

For the first part we have the following cases. Case INT follows by rule INT. Case VAR follows by Lemma 3.12. Case ABS1 follows by induction hypothesis and rule ABS1. Case ABS2 follows by induction hypothesis and rule ABS2. Case AABS1 follows by induction hypothesis and rule AABS1. Case AABS2 follows by the substitution property of $\vdash^{sub\sigma}$, induction hypothesis, and rule AABS2. Case APP follows by induction hypothesis, Lemma 3.12, and rule APP. Case ANNOT follows by induction hypothesis, Lemma 3.12, and rule ANNOT. Case LET follows by induction hypothesis and rule LET.

For the second part we have the following cases.

- Case GEN1. We have that $\Gamma \vdash^{poly}_{\Uparrow} t : \forall \overline{a}.\rho$, given that $\overline{a} = ftv(\rho) - ftv(\Gamma)$ and $\Gamma \vdash t : \rho$. Assume without loss of generality that $\overline{a} \notin vars(S)$. Consider a substitution $[\overline{a \mapsto b}]$ such that $\overline{b} \notin vars(S), ftv(\Gamma, \rho)$. By induction hypothesis $S[\overline{a \mapsto b}]\Gamma \vdash_{\Uparrow} t : S[\overline{a \mapsto b}]\rho$, or equivalently $S\Gamma \vdash_{\Uparrow} t : S[\overline{a \mapsto b}]\rho$.

  Now we can show that $\overline{b} = ftv([\overline{a \mapsto b}]S\rho) - ftv(S\Gamma)$. Suppose instead that there is a $b \notin ftv([\overline{a \mapsto b}]S\rho) - ftv(S\Gamma)$ which means that $b \in ftv(S\Gamma)$, since we know that $b \in ftv([\overline{a \mapsto b}]S\rho)$. But then, since $b \notin vars(S)$, it must be that $b \in ftv(\Gamma)$, a contradiction. On the other hand, suppose that we have a variable $g \in ftv([\overline{a \mapsto b}]S\rho) - ftv(S\Gamma)$ but $g \notin \overline{b}$. Note that it must be that $g \notin \overline{a}$ as well simply because $\overline{a} \notin ftv([\overline{a \mapsto b}]S\rho)$. Then this means that there is some variable $d \in \rho$ such that $g \in ftv(Sd)$. Now, d must be in $ftv(\Gamma)$, otherwise $d \in \overline{a}$ and then $g \in \overline{a}$. But if $d \in ftv(\Gamma)$ then $g \in ftv(\Gamma)$, a contradiction. Therefore it is indeed the case that $\overline{b} = ftv([\overline{a \mapsto b}]S\rho) - ftv(S\Gamma)$ and and we can apply the rule GEN1 to get the result.

- Case GEN2. In this case $\Gamma \vdash^{poly}_{\Downarrow} t : \forall \overline{a}.\rho$, given that $\overline{a} \notin ftv(\Gamma)$ and $\Gamma \vdash_{\Downarrow} t : \rho$. Consider a substitution $[\overline{a \mapsto c}]$ such that $\overline{c} \notin ftv(\Gamma), vars(S)$. Then by induction hypothesis $S[\overline{a \mapsto c}]\Gamma \vdash_{\Downarrow} t : S[\overline{a \mapsto c}]\rho$, or equivalently, $S\Gamma \vdash_{\Downarrow} t : S[\overline{a \mapsto c}]\rho$. Applying rule GEN2 we get $S\Gamma \vdash^{poly}_{\Downarrow} t : \forall \overline{c}.S[\overline{a \mapsto c}]\rho$, or $S\Gamma \vdash^{poly}_{\Downarrow} t : S(\forall \overline{a}.\rho)$.

$\square$

3.2.1 Shallow subsumption

Definition 3.14 (Shallow subsumption). *We define a subset of the subsumption relation, $\vdash^{sh} \cdot \leq \cdot$, which we call* shallow subsumption *as follows:*

$$\frac{\overline{b} \notin ftv(\forall \overline{a}.\rho)}{\vdash^{sh} \forall \overline{a}.\rho \leq \forall \overline{b}.[\overline{a \mapsto \tau}]\rho} \text{ SUB}$$

Notice that shallow subsumption is essentially ML subsumption. The rule SUB is equivalent to the rule SUB of the predicative F-eta subsumption.

Lemma 3.15. *If $\vdash^{sh} \sigma_1 \leq \sigma_2$ then $\vdash^{ol} \sigma_1 \leq \sigma_2$ and $\vdash^{\eth\mathfrak{s}\mathfrak{k}} \sigma_1 \leq \sigma_2$.*

Proof. The first part follows by application of Lemma 2.2 (reflexivity), rule SPEC, and rule SKOL. For the second part, $\vdash^{\eta} \sigma_1 \leq \sigma_2$ by rule SUB, and by Corollary 2.31 we get $\vdash^{\eth\mathfrak{s}\mathfrak{k}} \sigma_1 \leq \sigma_2$. $\square$

Lemma 3.16. *If $\vdash^{sh} \sigma_1 \leq \sigma_2$ then $ftv(\sigma_1) \subseteq ftv(\sigma_2)$.*

Proof. It must be that $\sigma_1 = \forall \overline{a}.\rho$ and $\sigma_2 = \forall \overline{b}.[\overline{a \mapsto \tau}]\rho$. Then for every $c \in ftv(\forall \overline{a}.\rho) = ftv(\rho) - \overline{a}$ it must be that $c \in ftv([\overline{a \mapsto \tau}]\rho)$. $\square$

Lemma 3.17. $\vdash^{sh} S\overline{\Gamma}(\rho) \leq \overline{S\Gamma}(S\rho)$.

Proof. Let $\overline{\Gamma}(\rho) = \forall \overline{a}.\rho$ where $\overline{a} = ftv(\rho) - ftv(\Gamma)$. Let $\overline{g}$ be a new set of variables, such that $\overline{g} \notin ftv(\Gamma), vars(S), ftv(\rho)$. Then $S\overline{\Gamma}(\rho) = \forall \overline{g}.S([\overline{a \mapsto g}]\rho)$.

Now, let $\overline{S\Gamma}(S\rho) = \forall \overline{b}.S\rho$, where $\overline{b} \in ftv(S\rho) - ftv(S\Gamma)$. We want to prove that $\vdash^{sh} \forall \overline{g}.S([\overline{a \mapsto g}]\rho) \leq \forall \overline{b}.S\rho$. First we need to show that $\overline{b} \notin ftv(S\overline{\Gamma}(\rho))$. By contradiction, assume that there exists a $b \in \overline{b}$ such that $b \in ftv(S\overline{\Gamma}(\rho))$. Therefore there exists $d \in ftv(\overline{\Gamma}(\rho))$ such that $b \in Sd$. From this we get that $d \in ftv(\rho)$ and $d \in ftv(\Gamma)$. Then, since $b \in Sd$, $b \in S\Gamma$, which is a contradiction to the fact that $b \in ftv(S\rho) - ftv(S\Gamma)$. Therefore, it only remains to be shown that for some types $\overline{\tau}$ it is the case that $[\overline{g \mapsto \tau}]S([\overline{a \mapsto g}]\rho) = S\rho$. Pick $\overline{\tau} = \overline{Sa}$. $\qquad \square$

Lemma 3.18 (Shallow Subsumption Weakening). *When $\vdash^{sub\sigma}$ is either $\vdash^{ol}$ or $\vdash^{\mathfrak{O}s\ell}$ the following are true:*

1. *If $\vdash^{sh} \sigma_1 \leq \sigma_2$ and $\vdash^{inst}_{\Downarrow} \sigma_2 \leq \rho$ then $\vdash^{inst}_{\Downarrow} \sigma_1 \leq \rho$.*

2. *If $\vdash^{sh} \sigma_1 \leq \sigma_2$ and $\vdash^{inst}_{\Uparrow} \sigma_2 \leq \rho$ then $\vdash^{inst}_{\Uparrow} \sigma_1 \leq \rho$.*

Moreover in each case the two derivations have the same height.

Proof. For the first part, by Lemma 3.15, $\vdash^{sub\sigma} \sigma_1 \leq \sigma_2$. By inversion $\vdash^{sub\sigma} \sigma_2 \leq \rho$, therefore by transitivity $\vdash^{sub\sigma} \sigma_1 \leq \rho$ and by rule INST2 we get the result.

For the second part, let $\sigma_1 = \forall \overline{a}.\rho_1$ and $\sigma_2 = \forall \overline{b}.[\overline{a \mapsto \tau_a}]\rho_1$, where $\overline{b} \notin ftv(\sigma_1)$. By inversion $\rho = [\overline{b \mapsto \tau_b}][\overline{a \mapsto \tau_a}]\rho_1$, or, since $\overline{b} \notin ftv(\sigma_1)$, $\rho = [\overline{a \mapsto S\tau_a}]\rho_1$, where $S = [\overline{b \mapsto \tau_b}]$. We get the result by applying rule INST1. $\qquad \square$

Lemma 3.19 (Weakening). *Given two contexts, Γ, Γ', if $dom(\Gamma) = dom(\Gamma')$ and for all $x \in dom(\Gamma)$ it is $\vdash^{sh} \Gamma(x) \leq \Gamma'(x)$ then the following are true:*

1. *If $\Gamma' \vdash_{\Uparrow} t : \rho$ then $\Gamma \vdash_{\Uparrow} t : \rho$.*

2. *If $\Gamma' \vdash_{\Downarrow} t : \rho$ then $\Gamma \vdash_{\Downarrow} t : \rho$.*

3. *If $\Gamma' \vdash^{poly}_{\Downarrow} t : \sigma$ then $\Gamma \vdash^{poly}_{\Downarrow} t : \sigma$.*

4. *If $\Gamma' \vdash^{poly}_{\Uparrow} t : \sigma$ then $\Gamma \vdash^{poly}_{\Uparrow} t : \sigma'$ where $\vdash^{sh} \sigma' \leq \sigma$.*

Moreover, for each implication, the two derivations have the same height.

Proof. We prove the four goals simultaneously by induction on the height of the derivations. For each goal the induction hypothesis asserts all others for any derivations of smaller height. We proceed by case analysis on the last rule used.

For the first goal we have the following cases for the last rule used in the derivation of $\Gamma' \vdash_{\Uparrow} t : \rho$.

- Case VAR. We have that $\Gamma' \vdash_{\Uparrow} x : \rho$, given that $\vdash^{inst}_{\Uparrow} \sigma \leq \rho$ and $x : \sigma \in \Gamma'$. By our assumptions, there exists a σ_0 such that $x : \sigma_0 \in \Gamma$ and $\vdash^{sh} \sigma_0 \leq \sigma$. Then the result follows from Lemma 3.18 and by applying rule VAR again.

- Case ABS1. Here $\Gamma' \vdash_{\Uparrow} \backslash x.t : \tau \rightarrow \rho$ given that $\Gamma', x : \tau \vdash_{\Uparrow} t : \rho$. By induction hypothesis $\Gamma, x : \tau \vdash_{\Uparrow} t : \rho$ and by the rule ABS1 we are done.

- Case AABS1. We have that $\Gamma' \vdash_\Uparrow \backslash x :: \sigma . t : \sigma \to \rho$ given that $\Gamma', x : \sigma \vdash_\Uparrow t : \rho$. By induction hypothesis $\Gamma, x : \sigma \vdash_\Uparrow t : \rho$ and by applying rule AABS1 we are done.

- Case APP. Here $\Gamma' \vdash_\Uparrow t\, u : \rho$ given that $\Gamma' \vdash_\Uparrow t : \sigma \to \sigma'$, $\Gamma' \vdash^{poly}_\Downarrow u : \sigma$ and $\vdash^{inst}_\Uparrow \sigma' \le \rho$. By induction hypothesis $\Gamma \vdash_\Uparrow t : \sigma \to \sigma'$ and $\Gamma \vdash^{poly}_\Downarrow u : \sigma$, and by applying rule APP we are done.

- Case LET. Here $\Gamma' \vdash_\Uparrow \mathtt{let}\ x\ =\ u\ \mathtt{in}\ t : \rho$ given that $\Gamma' \vdash^{poly}_\Uparrow u : \sigma$ and $\Gamma', x : \sigma \vdash_\Uparrow t : \rho$. By induction hypothesis $\Gamma \vdash^{poly}_\Uparrow u : \sigma'$ so that $\vdash^{sh} \sigma' \le \sigma$. Then, by induction hypothesis $\Gamma, x : \sigma' \vdash_\Uparrow t : \rho$ and by applying the rule LET we are done.

For the second goal we have the following cases for the last rule used in the derivation of $\Gamma' \vdash_\Downarrow t : \rho$.

- Case VAR. Similar to the case for VAR above.

- Case ABS2. We have that $\Gamma' \vdash_\Downarrow (\backslash x . t) : \sigma_a \to \sigma_r$, given that $\Gamma', x : \sigma_a \vdash^{poly}_\Downarrow t : \sigma_r$. By induction hypothesis $\Gamma, x : \sigma_a \vdash^{poly}_\Downarrow t : \sigma_r$, and by applying rule ABS2 this case is done.

- Case AABS2. Here $\Gamma' \vdash_\Downarrow \backslash x :: \sigma_x . t : \sigma_a \to \sigma_r$ given that $\vdash^{ol} \sigma_a \le \sigma_x$ and $\Gamma', (x : \sigma_x) \vdash^{poly}_\Downarrow t : \sigma_r$. Then, by induction hypothesis we get that $\Gamma, (x : \sigma_x) \vdash^{poly}_\Downarrow t : \sigma_r$ and by rule AABS2 we are done.

- Case APP. Similar to the case for APP above.

- Case LET. Similar to the case for LET above.

For the third part, $\Gamma' \vdash^{poly}_\Downarrow t : \sigma$ can be derived using the GEN2 rule. Let $\sigma = \forall \overline{a} . \rho$ and then we have that $\Gamma' \vdash^{poly}_\Downarrow t : \forall \overline{a} . \rho$ given that $\Gamma' \vdash_\Downarrow t : \rho$ for $\overline{a} \notin ftv(\Gamma')$. By induction hypothesis we get that $\Gamma \vdash_\Downarrow t : \rho$. Moreover, since the two contexts are related pointwise in the shallow subsumption relation, by Lemma 3.16 we have that $ftv(\Gamma) \subseteq ftv(\Gamma')$ hence $\overline{b} \notin ftv(\Gamma)$, and we can apply GEN2 to get the result.

Finally $\Gamma' \vdash^{poly}_\Uparrow t : \sigma$ is derivable using the GEN1 rule. Assume $\sigma = \forall \overline{a} . \rho$ where $\Gamma' \vdash_\Uparrow t : \rho$ and $\overline{a} = ftv(\rho) - ftv(\Gamma')$. By induction hypothesis $\Gamma \vdash_\Uparrow t : \rho$ and because $ftv(\Gamma) \subseteq ftv(\Gamma')$, using Lemma 3.16 it must be that $ftv(\rho) - ftv(\Gamma') \subseteq ftv(\rho) - ftv(\Gamma)$, which means that $\overline{a} \subseteq \overline{c}$ where $\overline{c} = ftv(\rho) - ftv(\Gamma)$. By applying rule GEN1 we get the result and it is easy to confirm that $\vdash^{sh} \forall \overline{c} . \rho \le \forall \overline{a} . \rho$. $\qquad \square$

3.2.2 Connection of syntax-directed and bidirectional type system

Lemma 3.20. *Let* $\vdash^{sub\sigma}$ *be* $\vdash^{ol}$. *Then*

1. *if* $\Gamma \vdash_{sd} t : \rho$ *then* $\Gamma \vdash_\Uparrow \rho$.

2. *if* $\Gamma \vdash^{poly}_{sd} t : \sigma$ *then* $\Gamma \vdash^{poly}_\Uparrow \sigma$.

Proof. We prove the two claims simultaneously by induction on the height of the derivations. We proceed with case analysis on the last rule used.

- Case INT. Directly follows by rule INT.

- Case VAR. Directly follows by rule VAR.

- Case ABS. Follows by induction hypothesis and rule ABS.

39

- Case APP. We have that $\Gamma \vdash_{sd} t\ u : \rho$ given that

$$\Gamma \vdash_{sd} t : (\sigma \to \sigma') \tag{1}$$

$$\Gamma \vdash_{sd}^{poly} u : \sigma_1 \tag{2}$$

$$\vdash^{sub\sigma} \sigma_1 \leq \sigma \tag{3}$$

$$\vdash^{inst} \sigma' \leq \rho \tag{4}$$

By induction hypothesis for (1) we get $\Gamma \vdash_{\Uparrow} t : \sigma \to \sigma'$ and moreover $\Gamma \vdash_{\Uparrow}^{poly} t : \sigma_1$, which, by Lemma 3.11 gives $\Gamma \vdash_{\Downarrow}^{poly} t : \sigma_1$. By (3) and Lemma 3.9 $\Gamma \vdash_{\Downarrow}^{poly} t : \sigma$ and using (4) and APP we get the result.

- Case LET. In this case we have that $\Gamma \vdash \mathtt{let}\ x\ \mathtt{=}\ u\ \mathtt{in}\ t : \rho$, given that

$$\Gamma \vdash_{sd}^{poly} u : \sigma \tag{5}$$

$$\Gamma, x : \sigma \vdash_{sd}^{poly} t : \rho \tag{6}$$

Then by induction $\Gamma \vdash_{\Uparrow}^{poly} u : \sigma$ and $\Gamma, x : \sigma \vdash_{\Uparrow} t : \rho$. Applying rule LET finishes the case.

- Case ANNOT. We have that $\Gamma \vdash (t\mathtt{::}\sigma) : \rho$, when

$$\Gamma \vdash_{sd}^{poly} t : \sigma' \tag{7}$$

$$\vdash^{sub\sigma} \sigma' \leq \sigma \tag{8}$$

$$\vdash^{inst} \sigma \leq \rho \tag{9}$$

By induction hypothesis we get $\Gamma \vdash_{\Uparrow}^{poly} t : \sigma'$, which by Lemma 3.11 gives $\Gamma \vdash_{\Downarrow}^{poly} t : \sigma'$. From this, (8), and Lemma 3.9 it must be that $\Gamma \vdash_{\Downarrow}^{poly} t : \sigma$. From this and (9) we can apply rule ANNOT to get the result.

For the second part, the case for GEN follows directly by rule GEN in the bidirectional system. $\qquad\square$

If we replace the relation $\vdash^{sub\sigma}$ with $\vdash^{\partial s\ell}$ the above theorem becomes false. The intuition is that type annotations may induce some deep skolemisation subsumption that will succeed in the syntax-directed system since there we generalise more and fail in the bidirectional where we check more! For example, consider $\Gamma = x : \forall a.\,a \to b \to b \to c, u : \mathtt{Int}$ and $\Gamma \vdash_{sd}^{poly} ((x\ u)\mathtt{::}(\mathtt{Int} \to \forall c.\mathtt{Int} \to c)) : \mathtt{Int} \to \forall c.\mathtt{Int} \to c$ but it is not derivable that $\Gamma \vdash_{\Uparrow}^{poly} ((x\ u)\mathtt{::}(\mathtt{Int} \to \forall c.\mathtt{Int} \to c)) : \mathtt{Int} \to \forall c.\mathtt{Int} \to c$ —in fact it is not typable at all. Notice that if all annotations and types of binders in the context were in prenex form, then it is easy to see that it would never make a difference whether $\vdash^{sub\sigma}$ was $\vdash^{ol}$ or $\vdash^{\partial s\ell}$ and the theorem above would be true.

Naturally the other direction does not hold; the bidirectional system is more powerful than the simple syntax-directed system. As an example let $\sigma_{id} = \forall a.\,a \to a$, $\Gamma = g : (\sigma_{id} \to \sigma_{id}) \to \mathtt{Int}$ and consider inferring $\Gamma \vdash_{\Uparrow} g\ (\backslash x.x) : \mathtt{Int}$. This will type-check, as it is checkable that $\Gamma \vdash_{\Downarrow} \backslash x.x : \sigma_{id} \to \sigma_{id}$. Nevertheless it is not derivable that $\Gamma \vdash_{sd} g\ x : \mathtt{Int}$ as it will not be derivable that $\vdash^{sub\sigma} \forall a.\,a \to a \leq \sigma_{id} \to \sigma_{id}$.

3.3 Final version of the bidirectional system: deep skolemisation in polytype checking

The bidirectional system with $\vdash^{\partial s\ell}$ is a rich one but lacks two important properties, namely Lemma 3.20 becomes false when $\vdash^{sub\sigma}$ is $\vdash^{\partial s\ell}$ and so does Lemma 3.9 when the types and contexts are related by $\vdash^{\partial s\ell}$ instead of $\vdash^{ol}$.

Consider the alternative rule for $\vdash_{\Downarrow}^{poly}$ given below.

$$\frac{\begin{array}{c} pr(\sigma) = \forall \overline{a}.\rho \\ \overline{a} \notin ftv(\Gamma) \quad \Gamma \vdash_{\Downarrow} t : \rho \end{array}}{\Gamma \vdash_{\Downarrow}^{poly} t : \sigma} \text{ GEN2*}$$

Here's an important property of this system.

Lemma 3.21. *If* $\Gamma \vdash_{\Downarrow} t : \rho$ *and* $pr(\rho) = \forall \overline{a}.\rho_0$ *where* $\overline{a} \notin ftv(\Gamma)$ *then* $\Gamma \vdash_{\Downarrow} t : \rho_0$ *and the new derivation has the same height.*

Proof. By induction on the height of the derivation $\Gamma \vdash_{\Downarrow} t : \rho$. We proceed with case analysis on the last rule used.

- Case INT. Directly follows by rule INT.

- Case VAR. We have that $\Gamma \vdash_{\Downarrow} x : \rho$ given that $\vdash_{\Downarrow}^{inst} \sigma \leq \rho$ where $x : \sigma \in \Gamma$. Equivalently $\vdash^{\mathfrak{d}\mathfrak{s}\mathfrak{k}} \sigma \leq \rho$. Then it must be that $\vdash^{dsk} \sigma \leq \rho$ as well and there is a canonical derivation that uses SKOL at the end. This means that $\vdash^{dsk} \sigma \leq \rho_0$ and by applying rule VAR again we get the result.

- Case ABS2. We have that $\Gamma \vdash_{\Downarrow} (\backslash x.t) : (\sigma_a \rightarrow \sigma_r)$ given that $\Gamma, (x : \sigma_a) \vdash_{\Downarrow}^{poly} t : \sigma_r$. It is easy to see that $\Gamma, (x : \sigma_a) \vdash_{\Downarrow}^{poly} t : pr(\sigma_r)$ and has the same height, since $pr(\sigma_r) = pr(pr(\sigma_r))$. Then by inversion $pr(\sigma_r) = \forall \overline{a}.\rho_r$ with $\overline{a} \notin ftv(\sigma_a, \Gamma)$ $\Gamma, (x : \sigma_a) \vdash_{\Downarrow} t : \rho_r$ and by applying rule GEN2* $\Gamma, (x : \sigma_a) \vdash_{\Downarrow}^{poly} t : \rho_r$ as well. By rule ABS2 we get $\Gamma \vdash_{\Downarrow} t : \sigma_a \rightarrow \rho_r$ as required.

- Case AABS2. In this case $\Gamma \vdash_{\Downarrow} (\backslash(x::\sigma_x).t) : (\sigma_a \rightarrow \sigma_r)$ given that $\vdash^{sub\sigma} \sigma_a \leq \sigma_x$ and $\Gamma, (x : \sigma_x) \vdash_{\Downarrow}^{poly} t : \sigma_r$. With the same argument as in the case for ABS2 it must be that $\Gamma, (x : \sigma_x) \vdash_{\Downarrow}^{poly} t : \rho_r$ where $pr(\sigma_r) = \forall \overline{a}.\rho_r$ with $\overline{a} \notin ftv(\sigma_a, \Gamma)$. Moreover it is easy to check that $\vdash^{\mathfrak{d}\mathfrak{s}\mathfrak{k}} pr(\sigma_a) \leq \sigma_a$. Applying rule AABS2 again gives $\Gamma, x : \sigma_x \vdash_{\Downarrow} t : \sigma_a \rightarrow \rho_r$ and finishes the case.

- Case APP. Here we have that $\Gamma \vdash_{\Downarrow} t\ u : \rho$ given that

$$\Gamma \vdash_{\Uparrow} t : (\sigma \rightarrow \sigma') \tag{1}$$

$$\Gamma \vdash_{\Downarrow}^{poly} u : \sigma \tag{2}$$

$$\vdash_{\Downarrow}^{inst} \sigma' \leq \rho \tag{3}$$

From (3) $\vdash^{dsk} \sigma' \leq \rho$. Consider the canonical derivation that ends with SKOL. Then, assuming that $\overline{a} \notin ftv(\sigma')$ as well without loss of generality $\vdash^{dsk} \sigma' \leq \rho_0$. Applying rule APP again gives the result.

- Case LET. In this case $\Gamma \vdash_{\delta} \texttt{let } x = u \texttt{ in } t : \rho$ given that $\Gamma \vdash_{\Uparrow}^{poly} u : \sigma$ and $\Gamma, x : \sigma \vdash_{\delta} t : \rho$. Notice that since $\overline{a} \notin ftv(\Gamma)$ and $ftv(\sigma) \in ftv(\Gamma)$ by inversion, it must be that $\overline{a} \notin ftv(\Gamma, x : \sigma)$, therefore the case is done by application of the induction hypothesis and rule LET.

- Case ANNOT. We have that $\Gamma \vdash_{\Downarrow} (t::\sigma) : \rho$ given that

$$\Gamma \vdash_{\Downarrow}^{poly} t : \sigma \tag{4}$$

$$\vdash_{\Downarrow}^{inst} \sigma \leq \rho \tag{5}$$

With a similar argument as in the case for APP we get that $\vdash^{dsk} \sigma \leq \rho_0$ and applying rule ANNOT finishes the case.

□

Some important properties first that carry along from Section 3 and still hold for this variation of the bidirectional type system.

Lemma 3.22.

1. If $\Gamma \vdash_\Uparrow t : \rho$ then $\Gamma \vdash_\Downarrow t : \rho$.

2. If $\Gamma \vdash_\Uparrow^{poly} t : \sigma$ then $\Gamma \vdash_\Downarrow^{poly} t : \sigma$.

Proof. The proof of the first part is exactly like the proof of Lemma 3.11. For the second part we have that $\Gamma \vdash_\Uparrow^{poly} t : \sigma$ by rule GEN1 and assume that $\sigma = \forall \overline{a}.\rho$ where $\overline{a} \notin ftv(\Gamma)$. Then $\Gamma \vdash_\Uparrow t : \rho$ which by induction hypothesis gives us that $\Gamma \vdash_\Downarrow t : \rho$. Now let $pr(\rho) = \forall \overline{b}.\rho_0$ such that $\overline{b} \notin \overline{a}, ftv(\Gamma)$. By Lemma 3.21 $\Gamma \vdash_\Downarrow t : \rho_0$. But notice that $pr(\sigma) = \forall \overline{a}\overline{b}.\rho_0$. Hence we can apply rule GEN2* to get the result. □

Lemma 3.23 (Substitution).

1. If $\Gamma \vdash_\delta t : \rho$ then $S\Gamma \vdash_\delta t : S\rho$.

2. If $\Gamma \vdash_\delta^{poly} t : \sigma$ then $S\Gamma \vdash_\delta^{poly} t : S\sigma$.

Proof. Exactly like the proof of Lemma 3.13 except for the GEN2* case for the second subgoal. In that case we have that $\Gamma \vdash_\Downarrow^{poly} t : \sigma$ when $pr(\sigma) = \forall \overline{a}.\rho$ and $\overline{a} \notin ftv(\Gamma)$ and $\Gamma \vdash_\Downarrow t : \rho$. Then, consider a substitution $S \cdot [\overline{a \mapsto c}]$ such that $\overline{c} \notin vars(S), ftv(\Gamma, \rho)$. By induction hypothesis $S\Gamma \vdash_v t : S \cdot [\overline{a \mapsto c}]\rho$. But notice that $pr(S\sigma) = S(pr(\sigma)) = S(\forall \overline{a}.\rho)$. With a renaming $S(\forall \overline{a}.\rho) = \forall \overline{c}.S[\overline{a \mapsto c}]\rho$ and we can apply rule GEN2* to get the result. □

Lemma 3.24 (Weakening). *Given two contexts,* Γ, Γ', *if* $dom(\Gamma) = dom(\Gamma')$ *and for all* $x \in dom(\Gamma)$ *it is* $\vdash^{sh} \Gamma(x) \leq \Gamma'(x)$ *then the following are true:*

1. If $\Gamma' \vdash_\Uparrow t : \rho$ then $\Gamma \vdash_\Uparrow t : \rho$.

2. If $\Gamma' \vdash_\Downarrow t : \rho$ then $\Gamma \vdash_\Downarrow t : \rho$.

3. If $\Gamma' \vdash_\Downarrow^{poly} t : \sigma$ then $\Gamma \vdash_\Downarrow^{poly} t : \sigma$.

4. If $\Gamma' \vdash_\Uparrow^{poly} t : \sigma$ then $\Gamma \vdash_\Uparrow^{poly} t : \sigma'$ where $\vdash^{sh} \sigma' \leq \sigma$.

Moreover, for each implication, the two derivations have the same height.

Proof. The proof is the same essentially as in Lemma 3.19 except for the case for GEN2* in the third subgoal. In this case $\Gamma' \vdash_\Downarrow^{poly} t : \sigma$. Let $pr(\sigma) = \forall \overline{a}.\rho$ and then we have that $\Gamma' \vdash_\Downarrow^{poly} t : \sigma$ given that $\Gamma' \vdash_\Downarrow t : \rho$ for $\overline{a} \notin ftv(\Gamma')$. By induction hypothesis we get that $\Gamma \vdash_\Downarrow t : \rho$. Moreover, since the two contexts are related pointwise in the shallow subsumption relation, by Lemma 3.16 we have that $ftv(\Gamma) \subseteq ftv(\Gamma')$ hence $\overline{a} \notin ftv(\Gamma)$, and we can apply GEN2* to get the result. □

Now, even in this system we cannot arbitrarily strengthen contexts in the $\vdash^{dsl}$ relation and check the same ρ-type[7]. For example, consider $\Gamma_1 = x : \forall abc. a \to \mathtt{Int} \to b \to b \to c$ and $\Gamma_2 = x : \forall a. a \to \mathtt{Int} \to \forall b. b \to$

---

[7] This is somewhat ugly because it means that when a programmer writes her program and wants to revise the type annotations in the rest of the program; he should change the type annotations so that the new types are only $\vdash^{ol}$ more general and not $\vdash^{dsl}$.

$\forall c . b \to c$. Then it is the case that $\Gamma_2 \vdash_\Downarrow x\ \mathit{True} : \mathtt{Int} \to \forall b . b \to \forall c . b \to c$ but this is not derivable when Γ_2 is replaced with Γ_1, although Γ_1 is a more general context in the $\vdash^{\partial s \ell}$ way. However a very important slight variation of the weakening lemma holds.

Lemma 3.25. *Independently of whether $\vdash^{sub\sigma}$ is $\vdash^{ol}$ or $\vdash^{\partial s \ell}$, if $\vdash^{ol} \Gamma_1 \le \Gamma_2$ pointwise then the following are true:*

1. *If $\Gamma_2 \vdash_\Uparrow t : \rho_2$ then $\exists \rho_1 . \Gamma_1 \vdash_\Uparrow t : \rho_1$ and $\vdash^{ol} \rho_1 \le \rho_2$.*

2. *If $\Gamma_2 \vdash_\Downarrow t : \rho_1$ and $\vdash^{ol} \rho_1 \le \rho_2$ then $\Gamma_1 \vdash_\Downarrow t : \rho_2$.*

3. *If $\Gamma_2 \vdash^{poly}_\Uparrow t : \sigma_2$ then $\exists \sigma_1 . \Gamma_1 \vdash^{poly}_\Uparrow t : \sigma_1$ and $\vdash^{ol} \sigma_1 \le \sigma_2$.*

4. *If $\Gamma_2 \vdash^{poly}_\Downarrow t : \sigma_1$ and $\vdash^{\partial s \ell} \sigma_1 \le \sigma_2$ then $\Gamma_1 \vdash^{poly}_\Downarrow t : \sigma_2$. Notice the fourth claim that allows now for $\vdash^{\partial s \ell}$ weakening.*

Proof. The proof remains exactly the same as the proof of Lemma 3.9 except for the fourth part. Here we have that $\Gamma_2 \vdash^{poly}_\Downarrow t : \sigma_1$ given that

$$pr(\sigma_1) = \forall \overline{a} . \rho_1 \tag{1}$$
$$\overline{a} \notin \mathit{ftv}(\Gamma_2) \tag{2}$$
$$\Gamma_2 \vdash_\Downarrow t : \rho_1 \tag{3}$$

We also have that $\vdash^{\partial s \ell} \sigma_1 \le \sigma_2$ or equivalently $\vdash^{dsk} \sigma_1 \le \sigma_2$. Consider the canonical derivation of this that ends with SKOL. Let $pr(\sigma_2) = \forall \overline{b} . \rho_2$ and without loss of generality $\overline{b} \notin \mathit{ftv}(\Gamma_1), \mathit{ftv}(\forall \overline{a} . \rho_1)$. Then by inversion it must be that $\vdash^{dsk} \sigma_1 \le \rho_2$. But $\vdash^{dsk} pr(\sigma_1) \le \sigma_1$ therefore $\vdash^{dsk} \forall \overline{a} . \rho_1 \le \rho_2$, or since SKOL can only be trivially applied $\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \le \rho_2$. By (3) and the substitution lemma $\vdash_\Downarrow t : [\overline{a \mapsto \tau}]\rho_1$ with the same height. By induction hypothesis $\vdash_\Downarrow t : \rho_2$ and by applying rule GEN2* we are done. $\qquad \square$

Corollary 3.26. *If $\Gamma \vdash^{poly}_\Downarrow t : \sigma_1$ and $\vdash^{\partial s \ell} \sigma_1 \le \sigma_2$ then $\Gamma \vdash^{poly}_\Downarrow t : \sigma_2$.*

Proof. Special case of the fourth subclaim of Lemma 3.25. $\qquad \square$

Lemma 3.27.

1. *If $\Gamma \vdash_{sd} t : \rho$ then $\Gamma \vdash_\Uparrow \rho$.*

2. *If $\Gamma \vdash^{poly}_{sd} t : \sigma$ then $\Gamma \vdash^{poly}_\Uparrow \sigma$.*

Proof. Exactly like the proof of Lemma 3.20 but now appealing to Corollary 3.26 in the cases for ANNOT and APP. $\qquad \square$

Notice now that Lemma 3.27 is independent of whether we use $\vdash^{\partial s \ell}$ or $\vdash^{ol}$.

Lemma 3.28 (Weakening). *Let $\vdash^{sub\sigma}$ be $\vdash^{\partial s \ell}$. Suppose that $\vdash^{\partial s \ell} \Gamma_1 \le \Gamma_2$. Then*

1. *if $\Gamma_2 \vdash_\Uparrow t : \rho_2$ then $\exists \rho_1 . \Gamma_1 \vdash_\Uparrow t : \rho_1$ and $\vdash^{\partial s \ell} \overline{\Gamma_1}(\rho_1) \le \overline{\Gamma_2}(\rho_2)$.*

2. *if $\Gamma_2 \vdash_\Downarrow t : \rho_1$ and $pr(\rho_1) = \rho_1$ and $pr(\rho_2) = \rho_2$ and $\vdash^{\partial s \ell} \rho_1 \le \rho_2$ then $\Gamma_1 \vdash_\Downarrow t : \rho_2$.*

3. *if $\Gamma_2 \vdash^{poly}_\Uparrow t : \sigma_2$ then $\exists \sigma_1 . \Gamma_1 \vdash^{poly}_\Uparrow t : \sigma_1$ and $\vdash^{\partial s \ell} \sigma_1 \le \sigma_2$.*

4. *if $\Gamma_2 \vdash^{poly}_\Downarrow t : \sigma_1$ and $\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2$ then $\Gamma_1 \vdash^{poly}_\Downarrow t : \sigma_2$.*

Proof. We prove the four claims simultaneously by induction on the height of the derivations. For each claim we assume that all others hold for derivations of smaller height. We proceed by case analysis on the last rule used.

Part 1: We have the following cases.

- Case INT. Just pick Int as ρ_1 again.

- Case VAR. We have that $\Gamma_2 \vdash_\Uparrow x : \rho_2$ given that $x : \sigma_2 \in \Gamma_2$ and

$$\vdash^{inst}_\Uparrow \sigma_2 \leq \rho_2 \tag{1}$$

  Then, $x : \sigma_1 \in \Gamma_1$ such that

$$\vdash^{\partial s\ell} \sigma_1 \leq \sigma_2 \tag{2}$$

  Assume that $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2)$. Then by the substitution lemma we get $\Gamma_2 \vdash_\Uparrow x : [\overline{b \mapsto d}]\rho_2$ for some $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \rho_2)$. By (1) and the substitution lemma we get $\vdash^{inst}_\Uparrow \sigma_2 \leq [\overline{b \mapsto d}]\rho_2$ and by transitivity of $\vdash^{\partial s\ell}$

$$\vdash^{\partial s\ell} \sigma_1 \leq [\overline{b \mapsto d}]\rho_2 \tag{3}$$

  Moreover assume that $\sigma_1 = \forall \overline{a}.\rho_1$ and without loss of generality assume that $\overline{a} \notin ftv(\Gamma_1)$. Then $\vdash^{inst}_\Uparrow \sigma_1 \leq \rho_1$. By (3) we get that $\vdash^{\partial s\ell} \forall \overline{a}.\rho_1 \leq [\overline{b \mapsto d}]\rho_2$. Consider $\overline{a}' = ftv(\rho_1) - ftv(\Gamma_1)$. Then $\overline{a} \subseteq \overline{a}'$ and consequently $\vdash^{\partial s\ell} \forall \overline{a}'.\rho_1 \leq [\overline{b \mapsto d}]\rho_2$. Then it must be that $\overline{d} \notin ftv(\forall \overline{a}'.\rho_1)$ because otherwise $\overline{d} \in ftv(\Gamma_1)$. Then by SKOL admissibility $\vdash^{\partial s\ell} \forall \overline{a}'.\rho_1 \leq \forall \overline{d}.[\overline{b \mapsto d}]\rho_2$, or equivalently $\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_1) \leq \overline{\Gamma_2}(\rho_2)$.

- Case ABS. Here we have that $\Gamma_2 \vdash_\Uparrow (\backslash x.t) : (\tau \to \rho_2)$, given that

$$\Gamma_2, x : \tau \vdash_\Uparrow t : \rho_2 \tag{4}$$

  Consider $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2, \tau)$ and a renaming substitution $[\overline{b \mapsto d}]$ where $\overline{d} \notin ftv(\Gamma_1, \Gamma_2, \tau, \rho_2)$. Then by (4) and the substitution lemma we get $\Gamma_2, x : \tau \vdash_\Uparrow t : [\overline{b \mapsto d}]\rho_2$. By induction hypothesis there exists a ρ_1 such that $\Gamma_1, x : \tau \vdash_\Uparrow t : \rho_1$ and

$$\vdash^{\partial s\ell} \forall \overline{a}.\rho_1 \leq \forall \overline{d}.[\overline{b \mapsto d}]\rho_2 \tag{5}$$

  where $\overline{a} = ftv(\rho_1) - ftv(\Gamma_1, \tau)$. By the rule ABS we get that $\Gamma_1 \vdash_\Uparrow (\backslash x.t) : (\tau \to \rho_1)$. We wish to show that

$$\vdash^{\partial s\ell} \forall \overline{a}_1.\tau \to \rho_1 \leq \forall \overline{a}_2.\tau \to \rho_2 \tag{6}$$

  where $\overline{a}_1 = ftv(\rho_1, \tau) - ftv(\Gamma)$ and $\overline{a}_2 = ftv(\rho_2, \tau) - ftv(\Gamma)$. Notice that if $\overline{c} = ftv(\tau) - ftv(\Gamma)$ then $\overline{a}_1 = \overline{a}\,\overline{c}$ and $\overline{a}_2 = \overline{b}\,\overline{c}$. From (5), since by SUB $\vdash^{\partial s\ell} \forall \overline{d}.[\overline{b \mapsto d}]\rho_2 \leq [\overline{b \mapsto d}]\rho_2$, and by transitivity we get that

$$\vdash^{\partial s\ell} \forall \overline{a}.\rho_1 \leq [\overline{b \mapsto d}]\rho_2 \tag{7}$$

  Then, by rule FUN $\vdash^{\partial s\ell} \tau \to \forall \overline{a}.\rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$ and $\vdash^{\partial s\ell} \forall \overline{a}.\tau \to \rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$ by transitivity and rule DISTRIB. By SUB and transitivity $\vdash^{\partial s\ell} \forall \overline{a}\,\overline{c}.\tau \to \rho_1 \leq \tau \to [\overline{b \mapsto d}]\rho_2$. Now we claim that $\overline{c} \notin ftv(\forall \overline{a}\,\overline{c}.\tau \to \rho_1)$ and $\overline{d} \notin ftv(\forall \overline{a}\,\overline{c}.\tau \to \rho_1)$ similarly. The former because we quantified over them, the latter because the opposite would mean that $\overline{d} \in ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get that $\vdash^{\partial s\ell} \forall \overline{a}\,\overline{c}.\tau \to \rho_1 \leq \forall \overline{d}\,\overline{c}.\tau \to [\overline{b \mapsto d}]\rho_2$ and by an α-renaming of $\overline{d}$ to $\overline{b}$ we are done.

- Case AABS. Similar to the case for ABS.

- Case APP. In this case we have that $\Gamma_2 \vdash_{\Uparrow} t\ u : \rho_2$ given that

$$\Gamma_2 \vdash_{\Uparrow} t : (\sigma \to \sigma') \tag{8}$$

$$\Gamma_2 \vdash_{\Downarrow} u : \sigma \tag{9}$$

$$\vdash_{\Uparrow}^{inst} \sigma' \le \rho_2 \tag{10}$$

Consider a renaming substitution $\overline{g}_1 = ftv(\sigma, \sigma', \rho_2) - ftv(\Gamma_2)$ to fresh $\overline{g}_2$, such that $\overline{g}_2 \notin ftv(\Gamma_1)$. Then by the substitution lemma (8) becomes

$$\Gamma_2 \vdash_{\Uparrow} t : (\sigma_0 \to \sigma'_0) \tag{11}$$

where $\sigma_0 = [\overline{g_1 \mapsto g_2}]\sigma$ and $\sigma'_0 = [\overline{g_1 \mapsto g_2}]\sigma'$. By induction hypothesis on (11) we get that there exists a ρ_t with

$$\Gamma_1 \vdash_{\Uparrow} t : \rho_t \tag{12}$$

$$\vdash^{\partial s\mathfrak{k}} \overline{\Gamma_1}(\rho_t) \le \forall \overline{g}_2.\sigma_0 \to \sigma'_0 \tag{13}$$

Because of the choice of $\overline{g}_2$, from equation (13) we get

$$\vdash^{\partial s\mathfrak{k}} \overline{\Gamma_1}(\rho_t) \le \sigma_0 \to \sigma'_0 \tag{14}$$

There are two possible cases for ρ_t. It is either a type variable $a \notin ftv(\Gamma_1)$ or it will be an arrow type $\sigma_1^t \to \sigma_2^t$.

- Assume that $\rho_t = \sigma_1^t \to \sigma_2^t$ and let $\overline{a} = ftv(\rho_t) - ftv(\Gamma_1)$. Then by equation (14) and Corollary 2.20 we get:

$$\vdash^{ol} \forall \overline{a}\overline{b}.\mathfrak{pr}(\sigma_1^t) \to \rho_2^t \le \forall \overline{c}.\mathfrak{pr}(\sigma_0) \to \rho'_0 \tag{15}$$

where

$$\overline{b} \notin ftv(\sigma_1^t) \tag{16}$$

$$\overline{c} \notin ftv(\sigma_0, \Gamma_1, \Gamma_2, \sigma_1^t, \sigma_2^t) \tag{17}$$

$$\mathfrak{pr}(\sigma_2^t) = \forall \overline{b}.\rho_2^t \tag{18}$$

$$\mathfrak{pr}(\sigma'_0) = \forall \overline{c}.\rho'_0 \tag{19}$$

By (15) and (17) it must be that

$$\vdash^{ol} \forall \overline{a}\overline{b}.\mathfrak{pr}(\sigma_1^t) \to \rho_2^t \le \mathfrak{pr}(\sigma_0) \to \rho'_0 \qquad \text{(by inversion)}$$

$$\Rightarrow \quad \vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}](\mathfrak{pr}(\sigma_1^t) \to \rho_2^t) \le \mathfrak{pr}(\sigma_0) \to \rho'_0 \qquad \text{(by inversion)}$$

$$\Rightarrow \quad \vdash^{ol} [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \to [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \mathfrak{pr}(\sigma_0) \to \rho'_0$$

From the last equation, by inversion we get that

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \le [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \tag{20}$$

$$\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \rho'_0 \tag{21}$$

From Corollary (2.20) and (20)

$$\vdash^{\partial s\mathfrak{k}} \sigma_0 \le [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{22}$$

From (12) and the substitution lemma, we get

$$\Gamma_1 \vdash_{\Uparrow} t : [\overline{a \mapsto \tau_a}](\sigma_1^t \to \sigma_2^t) \tag{23}$$

45

By the substitution lemma for (9) we have that $\Gamma_2 \vdash_{\Downarrow} u : \sigma_0$ and by induction hypothesis and (22) we have

$$\Gamma_1 \vdash_{\Downarrow} u : [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{24}$$

Then, if $[\overline{a \mapsto \tau_a}]\sigma_2^t = \forall \overline{g}_3 . \rho_1$, where without loss of generality $\overline{g}_3 \notin ftv(\Gamma_1)$ we have that $\vdash_{\Uparrow}^{inst}$ $[\overline{a \mapsto \tau_a}]\sigma_2^t \le \rho_1$. We have all the premises of the rule APP and applying it gives us that $\Gamma_1 \vdash t\ u : \rho_1$. Then it is the case that $\overline{\Gamma_1}(\rho_1) = \overline{\Gamma_1}([\overline{a \mapsto \tau_a}]\sigma_2^t)$. By (21) we get that

$$
\begin{aligned}
&\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \le \rho_0' \\
\Rightarrow\ &\vdash^{ol} \forall \overline{b} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \rho_0' \\
\Rightarrow\ &\vdash^{ol} \forall \overline{db} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \rho_0'
\end{aligned}
$$

where $\overline{d} = ftv(\overline{\tau}_a, \rho_2^t) - ftv(\Gamma_1)$. But now we know that $\overline{c} \notin ftv(\forall \overline{db} . [\overline{a \mapsto \tau_a}]\rho_2^t)$, because it must be that $ftv(\forall \overline{db} . [\overline{a \to \tau_a}]\rho_2^t) \subseteq ftv(\Gamma_1)$, and by (17) $\overline{c} \notin ftv(\Gamma_1)$. Then we can apply rule SKOL to get that $\vdash^{ol} \forall \overline{db} . [\overline{a \mapsto \tau_a}]\rho_2^t \le \forall \overline{c} . \rho_0'$, and by Corollary 2.20 $\vdash^{\partial s\ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \sigma_0'$. By the substitution lemma for (10) we have $\vdash_{\Uparrow}^{inst} \sigma_0' \le [\overline{g_1 \to g_2}]\rho_2$ and by transitivity we have that $\vdash^{\partial s\ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le [\overline{g_1 \mapsto g_2}]\rho_2$. Now it cannot be that $\overline{g}_2 \in \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t$ because $\overline{g}_2 \notin ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get $\vdash^{\partial s\ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \forall \overline{g}_2 . [\overline{g_1 \mapsto g_2}]\rho_2$ or by dropping useless quantifiers and α-renaming $\vdash^{\partial s\ell} \forall \overline{d} . [\overline{a \mapsto \tau_a}]\sigma_2^t \le \overline{\Gamma_2}(\rho_2)$ as required.

– Assume that $\rho_t = a$ and let $a \notin ftv(\Gamma_1)$. Then by equation (14) and Corollary 2.20 we get:

$$\vdash^{ol} \forall a . a \le \forall \overline{c} . \mathfrak{pr}(\sigma_0) \to \rho_0' \tag{25}$$

where

$$\overline{c} \notin ftv(\sigma_0, \Gamma_1, \Gamma_2) \tag{26}$$
$$\mathfrak{pr}(\sigma_0') = \forall \overline{c} . \rho_0' \tag{27}$$

By (25) and (26) and inversion on $\vdash^{ol}$ it must be that

$$\vdash^{ol} \tau_1 \to \tau_2 \le \mathfrak{pr}(\sigma_0) \to \rho_0' \tag{28}$$

Now yet one more inversion gives

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \le \tau_1 \tag{29}$$
$$\vdash^{ol} \tau_2 \le \rho_0' \tag{30}$$

From Corollary (2.20) and (29)

$$\vdash^{\partial s\ell} \sigma_0 \le \tau_1 \tag{31}$$

From (12) and the substitution lemma, we get

$$\Gamma_1 \vdash_{\Uparrow} t : \tau_1 \to \tau_2 \tag{32}$$

By the substitution lemma for (9) we have that $\Gamma_2 \vdash_{\Downarrow} u : \sigma_0$ and by induction hypothesis and (31) we have

$$\Gamma_1 \vdash_{\Downarrow} u : \tau_1 \tag{33}$$

Then $\vdash_{\Uparrow}^{inst} \tau_2 \le \tau_2$. We have all the premises of the rule APP and applying it gives us that $\Gamma_1 \vdash t\ u : \tau_2$. By (30) we get that

$$\vdash^{ol} \forall \overline{d} . \tau_2 \le \rho_0' \tag{34}$$

where $\overline{d} = ftv(\tau_2) - ftv(\Gamma_1)$. But now we know that $\overline{c} \notin ftv(\forall \overline{d}.\tau_2)$, because it must be that $ftv(\forall \overline{db}.[\overline{a \mapsto \tau_a}]\rho_2^t) \subseteq ftv(\Gamma_1)$, and by (26) $\overline{c} \notin ftv(\Gamma_1)$. Then we can apply rule SKOL to get that $\vdash^{ol} \forall \overline{d}.\tau_2 \leq \forall \overline{c}.\rho_0'$, and by Corollary 2.20 $\vdash^{\partial s\ell} \forall \overline{d}.\tau_2 \leq \sigma_0'$. By the substitution lemma for (10) we have $\vdash_{\Uparrow}^{inst} \sigma_0' \leq [\overline{g_1 \mapsto g_2}]\rho_2$ and by transitivity we have that $\vdash^{\partial s\ell} \forall \overline{d}.\tau_2 \leq [\overline{g_1 \mapsto g_2}]\rho_2$. Now it cannot be that $\overline{g}_2 \in \forall \overline{d}.\tau_2$ because $\overline{g}_2 \notin ftv(\Gamma_1)$. Then we can apply SKOL admissibility to get $\vdash^{\partial s\ell} \forall \overline{d}.\tau_2 \leq \forall \overline{g}_2.[\overline{g_1 \mapsto g_2}]\rho_2$ or by dropping useless quantifiers and α-renaming $\vdash^{\partial s\ell} \forall \overline{d}.\tau_2 \leq \overline{\Gamma_2}(\rho_2)$ as required.

- Case LET. In this case we have that $\Gamma_2 \vdash_{\Uparrow} \text{let } x = u \text{ in } t : \rho_2$, given that

$$\Gamma_2 \vdash_{\Uparrow}^{poly} u : \sigma \tag{35}$$

$$\Gamma_2, x : \sigma \vdash_{\Uparrow} t : \rho_2 \tag{36}$$

By induction hypothesis for (35) $\Gamma_1 \vdash_{\Uparrow} u : \sigma'$ such that $\vdash^{\partial s\ell} \sigma' \leq \sigma$. By induction hypothesis for (36) we get $\Gamma_1, x : \sigma' \vdash_{\Uparrow} t : \rho_1$ such that $\vdash^{\partial s\ell} \overline{\Gamma_1, x : \sigma'}(\rho_1) \leq \overline{\Gamma_2, x : \sigma}(\rho_2)$ or since σ' is generalised over Γ_1 and σ is generalised over Γ_2 this becomes $\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_1) \leq \overline{\Gamma_2}(\rho_2)$ as required. Applying rule LET finishes the case.

- Case ANNOT. We have that $\Gamma_2 \vdash_{\Uparrow} (t :: \sigma) : \rho_2$, given that

$$\Gamma_2 \vdash_{\Downarrow} t : \sigma \tag{37}$$

$$\vdash_{\Uparrow}^{inst} \sigma \leq \rho_2 \tag{38}$$

By induction hypothesis for (37) $\Gamma_1 \vdash_{\Downarrow} t : \sigma$. Assume now that $\sigma = \forall \overline{a}.\rho_1$. and without loss of generality, $\overline{a} \notin ftv(\Gamma_1, \Gamma_2)$. Then $\vdash_{\Uparrow}^{inst} \sigma \leq \rho_1$. Moreover $\overline{\Gamma_1}\rho_1 = \sigma$, since type annotations are closed. Additionally $\rho_2 = [\overline{a \mapsto \tau}]\rho_1$ for some $\overline{\tau}$. By applying rule ANNOT we get that $\Gamma_1 \vdash_{\Uparrow} (t :: \sigma) : \rho_1$. We finally have to show that $\vdash^{\partial s\ell} \overline{\Gamma_1}\rho_1 \leq \overline{\Gamma_2}([\overline{a \mapsto \tau}]\rho_1)$. Since type annotations are closed this is equivalent to showing that $\vdash^{\partial s\ell} \sigma \leq \overline{\Gamma_2}[\overline{a \mapsto \tau}]\rho_1$, or by SKOL admissibility it is enough to show that $\vdash^{\partial s\ell} \sigma \leq [\overline{a \mapsto \tau}]\rho_1$, which follows by (38).

Part 2: For this part we have the cases below.

- Case INT. In this case it must be that $\rho_2 = \text{Int}$ as well and we are done by rule INT.

- Case VAR. We have that $\Gamma_2 \vdash_{\Downarrow} x : \rho_1$, given that $x : \sigma_2 \in \Gamma_2$ and $\vdash_{\Downarrow}^{inst} \sigma_2 \leq \rho_1$. It must be then that $x : \sigma_1 \in \Gamma_1$ and by transitivity of $\vdash^{\partial s\ell}$, $\vdash_{\Downarrow}^{inst} \sigma_1 \leq \rho_1$ and by one more use of transitivity $\vdash_{\Downarrow}^{inst} \sigma_1 \leq \rho_2$. Applying rule VAR finishes the case.

- Case ABS2. In this case $\Gamma_2 \vdash_{\Downarrow} (\backslash x . t) : (\sigma_a \rightarrow \sigma_r)$ given that $\Gamma_2, (x : \sigma_a) \vdash_{\Downarrow}^{poly} t : \sigma_r$. By assumptions $\vdash^{\partial s\ell} \sigma_a \rightarrow \sigma_r \leq \rho_2$ or equivalently,

$$\vdash^{dsk} \sigma_a \rightarrow \sigma_r \leq \rho_2 \tag{39}$$

By (39), and since ρ_2 is already in weak prenex form by assumptions the only rule applicable is FUN. Therefore by inversion it must be that $\rho_2 = \sigma_{21} \rightarrow \sigma_{22}$, such that $\vdash^{dsk} \sigma_{21} \leq \sigma_a$ and $\vdash^{dsk} \sigma_r \leq \sigma_{22}$. Also it must be that $pr(\sigma_r) = \sigma_r$ and $pr(\sigma_{22}) = \sigma_{22}$. Then, by induction hypothesis we get that $\Gamma_1, (x : \sigma_{21}) \vdash_{\Downarrow}^{poly} t : \sigma_{22}$ and by applying rule ABS2 we get the result.

- Case AABS2. Here $\Gamma_2 \vdash_{\Downarrow} (\backslash (x :: \sigma_x) . t) : (\sigma_a \rightarrow \sigma_r)$ where

$$\vdash^{sub\sigma} \sigma_a \leq \sigma_x \tag{40}$$

$$\Gamma_2, (x : \sigma_x) \vdash_{\Downarrow}^{poly} t : \sigma_r \tag{41}$$

By assumptions $\vdash^{\partial s\ell} \sigma_a \to \sigma_r \leq \rho_2$ or equivalently

$$\vdash^{dsk} \sigma_a \to \sigma_r \leq \rho_2 \tag{42}$$

By inversion on a canonical derivation of (42), it must be that $\rho_2 = \sigma_{21} \to \sigma_{22}$, such that $\vdash^{dsk} \sigma_{21} \leq \sigma_a$ and $\vdash^{dsk} \sigma_r \leq \sigma_{22}$. Also it must be that $pr(\sigma_r) = \sigma_r$ and $pr(\sigma_{22}) = \sigma_{22}$. By (40) and transitivity of $\vdash^{dsk}$ we get $\vdash^{dsk} \sigma_{21} \leq \sigma_x$. Moreover, by induction hypothesis for (41) $\Gamma_1, (x : \sigma_x) \vdash^{poly}_{\Downarrow} t : \sigma_{22}$ and the result follows by AABS2.

- Case APP. In this case we have that $\Gamma_2 \vdash_{\Downarrow} t\ u : \rho_1$ given that

$$\Gamma_2 \vdash_{\Uparrow} t : (\sigma \to \sigma') \tag{43}$$
$$\Gamma_2 \vdash_{\Downarrow} u : \sigma \tag{44}$$
$$\vdash^{inst}_{\Downarrow} \sigma' \leq \rho_1 \tag{45}$$

Consider a renaming substitution $\overline{g}_1 = ftv(\sigma, \sigma', \rho_1) - ftv(\Gamma_2)$ to fresh $\overline{g}_2$, such that $\overline{g}_2 \notin ftv(\Gamma_1)$. Then by the substitution lemma (43) becomes

$$\Gamma_2 \vdash_{\Uparrow} t : (\sigma_0 \to \sigma'_0) \tag{46}$$

where $\sigma_0 = [\overline{g_1 \mapsto g_2}]\sigma$ and $\sigma'_0 = [\overline{g_1 \mapsto g_2}]\sigma'$. By induction hypothesis on (46) we get that there exists a ρ_t with

$$\Gamma_1 \vdash_{\Uparrow} t : \rho_t \tag{47}$$
$$\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_t) \leq \overline{g}_2 . \sigma_0 \to \sigma'_0 \tag{48}$$

Because of the choice of $\overline{g}_2$, from equation (48) we get

$$\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_t) \leq \sigma_0 \to \sigma'_0 \tag{49}$$

Moreover if $\rho_1^0 = [\overline{g_1 \mapsto g_2}]\rho_1$, by (49) and (45) it must be that

$$\vdash^{\partial s\ell} \overline{\Gamma_1}(\rho_t) \leq \sigma_0 \to \rho_1^0 \tag{50}$$

There are two possible cases for ρ_t. It is either a type variable $a \notin ftv(\Gamma_1)$ or it will be an arrow type $\sigma_1^t \to \sigma_2^t$.

- Assume that $\rho_t = \sigma_1^t \to \sigma_2^t$ and let $\overline{a} = ftv(\rho_t) - ftv(\Gamma_1)$. Then by equation (49) and Corollary 2.20 we get:

$$\vdash^{ol} \forall \overline{a}\overline{b} . \mathfrak{pr}(\sigma_1^t) \to \rho_2^t \leq \mathfrak{pr}(\sigma_0) \to \rho_1^1 \tag{51}$$

where

$$\overline{b} \notin ftv(\sigma_1^t) \tag{52}$$
$$\mathfrak{pr}(\sigma_2^t) = \forall \overline{b} . \rho_2^t \tag{53}$$
$$\mathfrak{pr}(\rho_1^0) = \rho_1^1 \tag{54}$$

Notice that we used the fact that ρ_1^0 is **in weak prenex form** in equation (54). By (51) it must be that

$$
\begin{aligned}
&\vdash^{ol} \forall \overline{a}\overline{b} . \mathfrak{pr}(\sigma_1^t) \to \rho_2^t \leq \mathfrak{pr}(\sigma_0) \to \rho_1^1 && \text{(by inversion)} \\
\Rightarrow\ &\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}](\mathfrak{pr}(\sigma_1^t) \to \rho_2^t) \leq \mathfrak{pr}(\sigma_0) \to \rho_1^1 && \text{(by inversion)} \\
\Rightarrow\ &\vdash^{ol} [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \to [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \leq \mathfrak{pr}(\sigma_0) \to \rho_1^1
\end{aligned}
$$

48

From the last equation, by inversion we get that

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \leq [\overline{a \mapsto \tau_a}]\mathfrak{pr}(\sigma_1^t) \tag{55}$$

$$\vdash^{ol} [\overline{a \mapsto \tau_a}, \overline{b \mapsto \tau_b}]\rho_2^t \leq \rho_1^1 \tag{56}$$

From Corollary (2.20) and (55)

$$\vdash^{\partial s\ell} \sigma_0 \leq [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{57}$$

From (47) and the substitution lemma, we get

$$\Gamma_1 \vdash_\Uparrow t : [\overline{a \mapsto \tau_a}](\sigma_1^t \to \sigma_2^t) \tag{58}$$

By the substitution lemma for (44) we have that $\Gamma_2 \vdash_\Downarrow u : \sigma_0$ and by induction hypothesis and (57) we have

$$\Gamma_1 \vdash_\Downarrow u : [\overline{a \mapsto \tau_a}]\sigma_1^t \tag{59}$$

From (56) we have that $\vdash^{ol} \forall \overline{b}.[\overline{a \mapsto \tau_a}]\rho_2^t \leq \rho_1^1$ or by Corollary 2.20

$$\vdash^{\partial s\ell} [\overline{a \mapsto \tau_a}]\sigma_2^t \leq \rho_1^0 \tag{60}$$

Consider the substitution $V = [\overline{g_2 \mapsto g_1}]$; then the last equation and (58), (59), and (60) become:

$$V\Gamma_1 \vdash_\Uparrow t : V[\overline{a \mapsto \tau_a}](\sigma_1^t \to \sigma_2^t) \tag{61}$$

$$V\Gamma_1 \vdash_\Downarrow u : V[\overline{a \mapsto \tau_a}]\sigma_1^t \tag{62}$$

$$\vdash^{\partial s\ell} V[\overline{a \mapsto \tau_a}]\sigma_2^t \leq V\rho_1^0 \tag{63}$$

But $V\Gamma_1 = \Gamma_1$ and $V\rho_1^0 = \rho_1$, therefore by transitivity and (63) $\vdash^{\partial s\ell} V[\overline{a \mapsto \tau_a}]\sigma_2^t \leq \rho_2$. Then we can apply rule APP to get $\Gamma_1 \vdash_\Downarrow t\ u : \rho_2$.

– Assume that $\rho_t = a$ and let $a \notin ftv(\Gamma_1)$. Then by equation (14) and Corollary 2.20 we get:

$$\vdash^{ol} \forall a . a \leq \mathfrak{pr}(\sigma_0) \to \rho_1^0 \tag{64}$$

since $\mathfrak{pr}(\rho_1^0) = \rho_1^1$. By (64) and inversion on $\vdash^{ol}$ it must be that

$$\vdash^{ol} \tau_1 \to \tau_2 \leq \mathfrak{pr}(\sigma_0) \to \rho_1^1 \tag{65}$$

Now yet one more inversion gives

$$\vdash^{ol} \mathfrak{pr}(\sigma_0) \leq \tau_1 \tag{66}$$

$$\vdash^{ol} \tau_2 \leq \rho_1^1 \tag{67}$$

From Corollary (2.20) and (66)

$$\vdash^{\partial s\ell} \sigma_0 \leq \tau_1 \tag{68}$$

From (12) and the substitution lemma, we get

$$\Gamma_1 \vdash_\Uparrow t : \tau_1 \to \tau_2 \tag{69}$$

By the substitution lemma for (9) we have that $\Gamma_2 \vdash_\Downarrow u : \sigma_0$ and by induction hypothesis and (68) we have

$$\Gamma_1 \vdash_\Downarrow u : \tau_1 \tag{70}$$

From (67) and Corollary 2.20 we have that

$$\vdash^{\partial s\ell} \tau_2 \leq \rho_1^0 \tag{71}$$

Consider the substitution $V = [\overline{g_2 \mapsto g_1}]$; then the last equation and (69), (70), and (71) become:

$$V\Gamma_1 \vdash_\Uparrow t : V\tau_1 \to V\tau_2 \tag{72}$$

$$V\Gamma_1 \vdash_\Downarrow u : V\tau_1 \tag{73}$$

$$\vdash^{\partial s\ell} V\tau_2 \le V\rho_1^0 \tag{74}$$

But $V\Gamma_1 = \Gamma_1$ and $V\rho_1^0 = \rho_1$, therefore by transitivity and (74) $\vdash^{\partial s\ell} V\tau_2 \le \rho_2$. Then we can apply rule APP to get $\Gamma_1 \vdash_\Downarrow t\ u : \rho_2$.

- Case ANNOT. Here $\Gamma_2 \vdash_\Downarrow (t::\sigma) : \rho_1$ given that

$$\Gamma_2 \vdash_\Downarrow^{poly} t : \sigma \tag{75}$$

$$\vdash_\Downarrow^{inst} \sigma \le \rho_1 \tag{76}$$

By induction hypothesis for (75) $\Gamma_1 \vdash_\Downarrow^{poly} t : \sigma$ and by transitivity of $\vdash^{\partial s\ell}$ we obtain $\vdash_\Downarrow^{inst} \sigma \le \rho_2$. The case is done by applying rule ANNOT once again.

- Case LET. In this case we have $\Gamma_2 \vdash_\Downarrow \texttt{let } x \texttt{ = } u \texttt{ in } t : \rho_1$ when

$$\Gamma_2 \vdash_\Uparrow^{poly} u : \sigma \tag{77}$$

$$\Gamma_2, x : \sigma \vdash_\Downarrow t : \rho_1 \tag{78}$$

By induction hypothesis for (77) $\Gamma_1 \vdash_\Uparrow^{poly} u : \sigma'$ such that $\vdash^{\partial s\ell} \sigma' \le \sigma$. Then, by induction hypothesis for (78) $\Gamma_1, x : \sigma' \vdash_\Downarrow t : \rho_2$ and we are done with an application of rule LET.

Part 3: For this part, by rule GEN1 we have that $\sigma_2 = \forall\overline{b}.\rho_2$ such that $\overline{b} = ftv(\rho_2) - ftv(\Gamma_2)$ and $\Gamma_2 \vdash_\Uparrow t : \rho_2$. By induction hypothesis for some ρ_1 $\Gamma_1 \vdash_\Uparrow t : \rho_1$ and $\vdash^{\partial s\ell} \sigma_2 \le \forall\overline{a}.\rho_1$, where $\overline{a} = ftv(\rho_1) - ftv(\Gamma_1)$. Applying the rule GEN1 finishes the case since we get $\Gamma_1 \vdash_\Uparrow^{poly} t : \forall\overline{a}.\rho_1$.

Part 4: By rule GEN2* we have that $\Gamma_2 \vdash_\Downarrow^{poly} t : \sigma_1$ given that $pr(\sigma_1) = \forall\overline{a}.\rho_1$ and $\overline{a} \notin ftv(\Gamma_2)$ and

$$\Gamma_2 \vdash_\Downarrow t : \rho_1 \tag{79}$$

We also have that $\vdash^{\partial s\ell} \sigma_1 \le \sigma_2$, or $\vdash^{dsk} \sigma_1 \le \sigma_2$. We know that $\vdash^{dsk} pr(\sigma_1) \le \sigma_2$ and by transitivity of $\vdash^{dsk}$

$$\vdash^{dsk} \forall\overline{a}.\rho_1 \le \sigma_2 \tag{80}$$

The canonical derivation of (80) must have the rule SKOL applied last. Assume then that $pr(\sigma_2) = \forall\overline{b}.\rho_2$ and without loss of generality

$$\overline{b} \notin ftv(\Gamma_1, \sigma_1) \tag{81}$$

Then it must be that $\vdash^{dsk} \forall\overline{a}.\rho_1 \le \rho_2$ and by inversion $\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \le \rho_2$, for some $\overline{\tau}$. Moreover $pr([\overline{a \mapsto \tau}]\rho_1) = [\overline{a \mapsto \tau}]\rho_1$ and $pr(\rho_2) = \rho_2$. From (79) and the substitution lemma we get

$$\Gamma_2 \vdash_\Downarrow t : [\overline{a \mapsto \tau}]\rho_1 \tag{82}$$

From (82) and induction hypothesis we get $\Gamma_1 \vdash_\Downarrow t : \rho_2$ and because of (81) we can apply rule GEN2* to get the result. $\qquad\square$

$$\boxed{pr(\sigma) = \forall \overline{a}.\rho \mapsto t}$$

$$\frac{pr(\rho_1) = \forall \overline{b}.\rho_2 \mapsto t \quad \overline{a} \notin \overline{b}}{pr(\forall \overline{a}.\rho_1) = \forall \overline{a}\overline{b}.\rho_2 \mapsto \lambda x : (\forall \overline{a}\overline{b}.\rho_2).\Lambda \overline{a}.t \; (x \; [\overline{a}])} \; \text{PRPOLY}$$

$$\frac{pr(\sigma_2) = \forall \overline{a}.\rho_2 \mapsto t \quad \overline{a} \notin ftv(\sigma_1)}{pr(\sigma_1 \to \sigma_2) = \forall \overline{a}.\sigma_1 \to \rho_2 \mapsto \lambda x : (\forall \overline{a}.\sigma_1 \to \rho_2).\lambda y : \sigma_1.t \; (\Lambda \overline{a}.x \; [\overline{a}] \; y)} \; \text{PRFUN}$$

$$\frac{}{pr(\tau) = \tau \mapsto \lambda x : \tau.x} \; \text{PRMONO}$$

$$\boxed{\vdash^{dsk} \sigma \leq \sigma' \mapsto t}$$

$$\frac{\begin{array}{c} pr(\sigma_2) = \forall \overline{a}.\rho \mapsto t_1 \\ \overline{a} \notin ftv(\sigma_1) \quad \vdash^{dsk} \sigma_1 \leq \rho \mapsto t_2 \end{array}}{\vdash^{dsk} \sigma_1 \leq \sigma_2 \mapsto \lambda x : \sigma_1.t_1 \; (\Lambda \overline{a}.t_2 \; x)} \; \text{DEEP-SKOL}$$

$$\frac{\vdash^{dsk} [\overline{a \mapsto \tau}]\rho_1 \leq \rho_2 \mapsto t}{\vdash^{dsk} \forall \overline{a}.\rho_1 \leq \rho_2 \mapsto \lambda x : (\forall \overline{a}.\rho).t \; (x \; [\overline{\tau}])} \; \text{SPEC}$$

$$\frac{\vdash^{dsk} \sigma_3 \leq \sigma_1 \mapsto t_1 \qquad \vdash^{dsk} \sigma_2 \leq \sigma_4 \mapsto t_2}{\vdash^{dsk} (\sigma_1 \to \sigma_2) \leq (\sigma_3 \to \sigma_4) \mapsto \lambda x : \sigma_1 \to \sigma_2.\lambda y : \sigma_3.t_2 \; (x \; (t_1 \; y))} \; \text{FUN}$$

$$\frac{}{\vdash^{dsk} \tau \leq \tau \mapsto \lambda x : \tau.x} \; \text{MONO}$$

Figure 10: Creating coercion terms

$$\boxed{\Gamma \vdash^F t : \sigma}$$

$$\frac{x : \sigma \in \Gamma}{\Gamma \vdash^F x : \sigma} \; \text{VAR} \qquad \frac{}{\Gamma \vdash^F i : \texttt{Int}} \; \text{INT}$$

$$\frac{\Gamma, x : \sigma_1 \vdash^F t : \sigma_2}{\Gamma \vdash^F \lambda x : \sigma_1.t : \sigma_1 \to \sigma_2} \; \text{ABS} \qquad \frac{\Gamma \vdash^F t : \sigma_1 \to \sigma_2 \quad \Gamma \vdash^F u : \sigma_1}{\Gamma \vdash^F t \; u : \sigma_2} \; \text{APP}$$

$$\frac{\Gamma \vdash^F t : \sigma \quad a \notin ftv(\Gamma)}{\Gamma \vdash^F \Lambda a.t : \forall a.\sigma} \; \text{TABS} \qquad \frac{\Gamma \vdash^F t : \forall a.\sigma}{\Gamma \vdash^F t \; [\sigma_1] : [a \mapsto \sigma_1]\sigma} \; \text{TAPP}$$

Figure 11: System-F with open types

$$\boxed{\Gamma \vdash_\delta t_1 : \rho \mapsto t_2}$$

$$\frac{}{\Gamma \vdash_\delta i : \mathtt{Int} \mapsto i} \text{ INT} \qquad \frac{\vdash_\delta^{inst} \sigma \le \rho \mapsto f}{\Gamma, (x : \sigma) \vdash_\delta x : \rho \mapsto f\ x} \text{ VAR}$$

$$\frac{\Gamma, (x : \tau) \vdash_\Uparrow t : \rho \mapsto t_1}{\Gamma \vdash_\Uparrow (\backslash x\,.\,t) : (\tau \to \rho) \mapsto \lambda x{:}\tau\,.\,t_1} \text{ ABS1}$$

$$\frac{\Gamma, (x : \sigma_a) \vdash_\Downarrow^{poly} t : \sigma_r \mapsto t_1}{\Gamma \vdash_\Downarrow (\backslash x\,.\,t) : (\sigma_a \to \sigma_r) \mapsto \lambda x{:}\sigma_a\,.\,t_1} \text{ ABS2}$$

$$\frac{\Gamma, (x : \sigma) \vdash_\Uparrow t : \rho \mapsto t_1}{\Gamma \vdash_\Uparrow (\backslash(x{::}\sigma)\,.\,t) : (\sigma \to \rho) \mapsto \lambda x{:}\sigma\,.\,t_1} \text{ AABS1}$$

$$\frac{\vdash^{dsk} \sigma_a \le \sigma_x \mapsto f \qquad \Gamma, (x : \sigma_x) \vdash_\Downarrow^{poly} t : \sigma_r \mapsto t_1}{\Gamma \vdash_\Downarrow (\backslash(x{::}\sigma_x)\,.\,t) : (\sigma_a \to \sigma_r) \mapsto \lambda x{:}\sigma_a\,.\,[x \mapsto (f\ x)]t_1} \text{ AABS2}$$

$$\frac{\Gamma \vdash_\Uparrow t : (\sigma \to \sigma') \mapsto t_1 \quad \Gamma \vdash_\Downarrow^{poly} u : \sigma \mapsto u_1 \quad \vdash_\delta^{inst} \sigma' \le \rho \mapsto f}{\Gamma \vdash_\delta t\ u : \rho \mapsto f\ (t_1\ u_1)} \text{ APP}$$

$$\frac{\begin{array}{c}\Gamma \vdash_\Downarrow^{poly} t : \sigma \mapsto t_1 \\ \vdash_\delta^{inst} \sigma \le \rho \mapsto f\end{array}}{\Gamma \vdash_\delta (t{::}\sigma) : \rho \mapsto f\ t_1} \text{ ANNOT} \qquad \frac{\begin{array}{c}\Gamma \vdash_\Uparrow^{poly} u : \sigma \mapsto u_1 \\ \Gamma, x : \sigma \vdash_\delta t : \rho \mapsto t_1\end{array}}{\Gamma \vdash_\delta \mathtt{let}\ x\ =\ u\ \mathtt{in}\ t : \rho \mapsto (\lambda x{:}\sigma\,.\,t_1)\ u_1} \text{ LET}$$

$$\boxed{\Gamma \vdash_\delta^{poly} t : \sigma \mapsto t_1}$$

$$\frac{\begin{array}{c}\overline{a} = ftv(\rho) - ftv(\Gamma) \\ \Gamma \vdash_\Uparrow t : \rho \mapsto t_1\end{array}}{\Gamma \vdash_\Uparrow^{poly} t : \forall \overline{a}\,.\,\rho \mapsto \Lambda \overline{a}\,.\,t_1} \text{ GEN1} \qquad \frac{\begin{array}{c}pr(\sigma) = \forall \overline{a}\,.\,\rho \mapsto f \\ \overline{a} \notin ftv(\Gamma) \quad \Gamma \vdash_\Downarrow t : \rho \mapsto t_1\end{array}}{\Gamma \vdash_\Downarrow^{poly} t : \sigma \mapsto f\ (\Lambda \overline{a}\,.\,t_1)} \text{ GEN2}$$

$$\boxed{\vdash_\delta^{inst} \sigma \le \rho \mapsto t}$$

$$\frac{}{\vdash_\Uparrow^{inst} \forall \overline{a}\,.\,\rho \le [\overline{a} \mapsto \overline{\tau}]\,\rho \mapsto \lambda x{:}(\forall \overline{a}\,.\,\rho)\,.\,x[\overline{\tau}]} \text{ INST1} \qquad \frac{\vdash^{dsk} \sigma \le \rho \mapsto t}{\vdash_\Downarrow^{inst} \sigma \le \rho \mapsto t} \text{ INST2}$$

Figure 12: Bidirectional higher-rank type system with retyping functions

3.3.1 Type-safety of the bidirectional system

The semantics of the language is defined via a translation to System-F terms (where open types are allowed and treated as arbitrary constants). Subsumption creates coercion terms that are applied appropriately. We give an more suitable presentation of weak prenex conversion and the subsumption relation in Figure 10. System-F semantics are given in Figure 11. The actual translation is given in Figure 12.

Lemma 3.29 (Weak prenex retyping). *If $pr(\sigma) = \forall \overline{a}.\rho \mapsto t$ then $\vdash^F t : (\forall \overline{a}.\rho) \to \sigma$.*

Proof. Easy induction. $\qquad \square$

Lemma 3.30 (Subsumption retyping). *If $\vdash^{dsk} \sigma_1 \leq \sigma_2 \mapsto t$ then $\vdash^F t : \sigma_1 \to \sigma_2$.*

Proof. Easy induction. $\qquad \square$

Lemma 3.31 (Translation semantics).

1. *If $\Gamma \vdash_\delta t_1 : \rho \mapsto t_2$ then $\Gamma \vdash^F t_2 : \rho$.*

2. *If $\Gamma \vdash_\delta^{poly} t_1 : \sigma \mapsto t_2$ then $\Gamma \vdash^F t_2 : \sigma$.*

Proof. Easy induction. $\qquad \square$

Corollary 3.32 (Type soundness). *The bidirectional system has the type soundness property.*

Proof. By Lemma 3.31 the translation yields a well-typed System-F term. $\qquad \square$

3.4 Conservativity over Damas-Milner

We show that the type systems we introduced are all conservative extensions of the Damas-Milner type system, given in Figure 13 and Figure 14. Damas-Milner types are of the form $\forall \overline{a}.\tau$.

Definition 3.33 ($\mathcal{DM}(\Gamma, t)$ predicate). *Let the predicate $\mathcal{DM}(\Gamma, t)$ where t is a term and Γ a context be true iff:*

- *All types bound in Γ are DM types.*

- *t contains no type annotations.*

Lemma 3.34. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash^{DM} t : \sigma$ then $\Gamma \vdash_{nsd} t : \sigma$.*

Proof. The proof is by induction on the height of the derivation $\Gamma \vdash^{DM} t : \sigma$ which is completely straightforward and we ommit. $\qquad \square$

However it is not true that if $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{nsd} t : \sigma$ then $\Gamma \vdash^{DM} t : \sigma$, since by rule SUBS we can downgrade the type arbitrarily. What is true is that $\exists \sigma'.\Gamma \vdash^{DM} t : \sigma'$ such that $\vdash^{dsk} \sigma' \leq \sigma$.

Lemma 3.35. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{sd} t : \tau$ then $\Gamma \vdash_\Uparrow t : \tau$.*

Proof. Consequence of Lemma 3.27. $\qquad \square$

Lemma 3.36.

1. *If $\vdash^{ol} \forall \overline{a}.\tau_1 \leq \tau_2$ then $[\overline{a \mapsto \tau}]\tau_1 = \tau_2$ for some $\overline{\tau}$.*

$$\text{Rho-types} \quad \rho \quad ::= \quad \tau$$

$$\boxed{\Gamma \vdash t : \sigma}$$

$$\frac{}{\Gamma \vdash i : \texttt{Int}} \text{ INT} \qquad \frac{}{\Gamma, (x : \sigma) \vdash x : \sigma} \text{ VAR}$$

$$\frac{\Gamma, (x : \tau) \vdash t : \rho}{\Gamma \vdash (\backslash x \,.\, t) : (\tau \to \rho)} \text{ ABS} \qquad \frac{\Gamma \vdash t : \tau \to \rho \quad \Gamma \vdash u : \tau}{\Gamma \vdash t \; u : \rho} \text{ APP}$$

$$\frac{\Gamma \vdash u : \sigma \quad \Gamma, x : \sigma \vdash t : \rho}{\Gamma \vdash \texttt{let } x \texttt{ = } u \texttt{ in } t : \rho} \text{ LET} \qquad \frac{\Gamma \vdash t : \sigma}{\Gamma \vdash (t :: \sigma) : \sigma} \text{ ANNOT}$$

$$\frac{\overline{a} \notin \mathit{ftv}(\Gamma) \quad \Gamma \vdash t : \rho}{\Gamma \vdash t : \forall \overline{a} . \rho} \text{ GEN} \qquad \frac{\Gamma \vdash t : \forall \overline{a} . \rho}{\Gamma \vdash t : [\overline{a} \mapsto \overline{\tau}] \, \rho} \text{ INST}$$

Figure 13: The non-syntax-directed Damas-Milner type system

2. *If $\vdash^{\partial s \mathfrak{k}} \forall \overline{a} . \tau_1 \leq \tau_2$ then $[\overline{a \mapsto \tau}] \tau_1 = \tau_2$ for some $\overline{\tau}$.*

Proof. The first part follows by inversion on $\vdash^{ol}$. For the second, just observe that the prenex forms of the types are the types themselves, and by using Corollary 2.20 the result follows by the first part. $\qquad \square$

Lemma 3.37. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{\Uparrow} t : \rho$ then $\rho = \tau$.*

Proof. By induction on the derivation of $\Gamma \vdash_{\Uparrow} t : \rho$. None of the cases are interesting. $\qquad \square$

Lemma 3.38. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{sd} t : \rho$ then $\rho = \tau$.*

Proof. By induction on the derivation of $\Gamma \vdash_{\Uparrow} t : \rho$. None of the cases are interesting. $\qquad \square$

Lemma 3.39.

1. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{\Uparrow} t : \tau$ then $\Gamma \vdash^{DM}_{sd} t : \tau$.*

2. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{\Downarrow} t : \tau$ then $\Gamma \vdash^{DM}_{sd} t : \tau$.*

Proof. We prove the two claims simultaneously by induction on the height of the derivations. We proceed with case analysis on the last rule used.

- Case INT. Directly follows by INT.

- Case VAR. We have that $\Gamma \vdash_{\delta} x : \tau$, given that $\vdash^{inst}_{\delta} \sigma \leq \tau$, where $x : \sigma \in \Gamma$. In the $\Uparrow$ direction we can just apply VAR. In the $\Downarrow$ direction we have by assumptions that $\sigma = \forall \overline{a} . \tau_0$ and it is the case that $\vdash^{sub\sigma} \forall \overline{a} . \tau_0 \leq \tau$. By Lemma 3.36, $\tau = [\overline{a \mapsto \tau_2}] \tau_0$ and therefore $\vdash^{inst} \forall \overline{a} . \tau_0 \leq \tau$. We can then apply rule VAR to get the result.

Rho-types $\rho ::= \tau$

$$\boxed{\Gamma \vdash t : \rho}$$

$$\frac{}{\Gamma \vdash i : \texttt{Int}} \text{ INT} \qquad \frac{\vdash^{inst} \sigma \leq \rho}{\Gamma, (x : \sigma) \vdash x : \rho} \text{ VAR}$$

$$\frac{\Gamma, (x : \tau) \vdash t : \rho}{\Gamma \vdash (\backslash x \,.\, t) : (\tau \to \rho)} \text{ ABS} \qquad \frac{\Gamma \vdash t : \tau \to \rho \quad \Gamma \vdash u : \tau}{\Gamma \vdash t \ u : \rho} \text{ APP}$$

$$\frac{\Gamma \vdash^{poly} u : \sigma \quad \Gamma, x : \sigma \vdash t : \rho}{\Gamma \vdash \texttt{let } x \texttt{ = } u \texttt{ in } t : \rho} \text{ LET} \qquad \frac{\Gamma \vdash^{poly} t : \sigma' \quad \vdash^{sub\sigma} \sigma' \leq \sigma \quad \vdash^{inst} \sigma \leq \rho}{\Gamma \vdash (t::\sigma) : \rho} \text{ ANNOT}$$

$$\boxed{\Gamma \vdash^{poly} t : \sigma} \qquad\qquad \boxed{\vdash^{inst} \sigma \leq \rho}$$

$$\frac{\overline{a} = ftv(\rho) - ftv(\Gamma) \quad \Gamma \vdash t : \rho}{\Gamma \vdash^{poly} t : \forall \overline{a} \,.\, \rho} \text{ GEN} \qquad \frac{}{\vdash^{inst} \forall \overline{a} \,.\, \rho \ \leq \ [\overline{a \mapsto \tau}] \, \rho} \text{ INST}$$

$$\boxed{\vdash^{sub\sigma} \sigma \leq \sigma'}$$

$$\frac{\overline{a} \notin ftv(\sigma) \quad \vdash^{sub\sigma} \sigma \leq \rho}{\vdash^{sub\sigma} \sigma \leq \forall \overline{a} \,.\, \rho} \text{ SKOL} \qquad \frac{\vdash^{sub\sigma} [\overline{a \mapsto \tau}] \, \rho_1 \leq \rho_2}{\vdash^{sub\sigma} \forall \overline{a} \,.\, \rho_1 \leq \rho_2} \text{ SPEC} \qquad \frac{}{\vdash^{sub\sigma} \tau \leq \tau} \text{ MONO}$$

Figure 14: The syntax-directed Damas-Milner type system

- Case ABS1. Here we have that $\Gamma \vdash_\Uparrow \backslash x \,.\, t : \tau_1 \to \tau_2$ given that $\Gamma, x : \tau_1 \vdash_\Uparrow t : \tau_2$. By induction $\Gamma, x : \tau_1 \vdash^{DM}_{sd} t : \tau_2$ and applying rule ABS finishes the case.

- Case ABS2. $\Gamma \vdash_\Downarrow \backslash x \,.\, t : \tau_1 \to \tau_2$ given that $\Gamma, x : \tau_1 \vdash^{poly}_\Downarrow t : \tau_2$, or by inversion $\Gamma, x : \tau_1 \vdash_\Downarrow t : \tau_2$. By induction $\Gamma, x : \tau_1 \vdash^{DM}_{sd} t : \tau_2$ and applying rule ABS finishes the case.

- Case APP. We have that $\Gamma \vdash_\delta t \ u : \tau$ given that $\Gamma \vdash_\Uparrow t : \sigma \to \sigma'$, $\Gamma \vdash^{poly}_\Downarrow u : \sigma$ and $\vdash^{inst}_\delta \sigma' \leq \tau$. By Lemma 3.37 it must be that $\sigma = \tau_1 \to \tau_2$. Then by induction $\Gamma \vdash^{DM}_{sd} t : \tau_1 \to \tau_2$ and $\Gamma \vdash^{DM}_{sd} u : \tau_1$ and by APP $\Gamma \vdash^{DM}_{sd} t \ u : \tau_2$. But we know that $\vdash^{inst}_\delta \tau_2 \leq \tau$, hence $\vdash^{sub\sigma} \tau_2 \leq \tau$ and it can only be that $\tau_2 = \tau$.
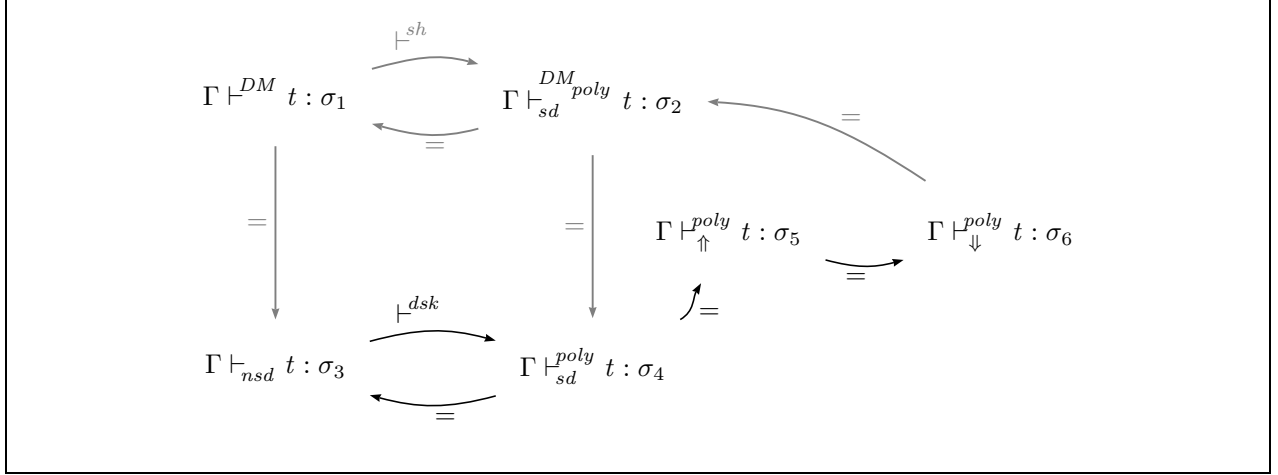
Figure 15: The world

- Case LET. In this case $\Gamma \vdash_\delta \text{let } x = u \text{ in } t : \tau$ given that

$$\Gamma \vdash^{poly}_\Uparrow u : \sigma \tag{1}$$

$$\Gamma, x : \sigma \vdash_\delta t : \tau \tag{2}$$

It must be the case that $\sigma = \forall \overline{a}.\rho$ such that $\Gamma \vdash_\Uparrow u : \rho$ and $\overline{a} = ftv(\rho) - ftv(\Gamma)$. By Lemma 3.37 we get that $\rho = \tau_1$ and by induction $\Gamma \vdash^{DM}_{sd} u : \tau_1$. By GEN then $\Gamma \vdash^{DM}_{sd}{}^{poly} u : \sigma$ and σ is an DM type. By induction hypothesis for (2) and application of LET we get the result.

The rest of the cases cannot happen. □

Lemma 3.40. *Assume that $\vdash^{sub\sigma}$ is $\vdash^{\partial s\ell}$ in what follows. If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash_{nsd} t : \sigma$ then there exists a σ' such that $\Gamma \vdash^{DM} t : \sigma'$ and $\vdash^{\partial s\ell} \sigma' \leq \sigma$.*

Proof. If $\Gamma \vdash_{nsd} t : \sigma$, by Lemma 3.4 we get $\Gamma \vdash_{sd} t : \rho$ such that $\vdash^{\partial s\ell} \overline{\Gamma}(\rho) \leq \sigma$. By Lemma 3.38 we get that $\rho = \tau$. Then by Lemma 3.35 we get that $\Gamma \vdash_\Uparrow t : \tau$ and by Lemma 3.39 we get $\Gamma \vdash^{DM}_{sd} t : \tau$. By Hindley and Milner's result $\Gamma \vdash^{DM} t : \tau$ and by rule GEN $\Gamma \vdash^{DM} t : \overline{\Gamma}(\tau)$. □

Lemma 3.41. *If $\mathcal{DM}(\Gamma, t)$ and $\Gamma \vdash^{DM}_{sd} t : \tau$ then $\Gamma \vdash_{sd} t : \tau$ and if $\Gamma \vdash^{DM}_{sd}{}^{poly} t : \sigma$ then $\Gamma \vdash^{poly}_{sd} t : \sigma$.*

Proof. Straightforward induction. □

In conclusion, the world looks like Figure 15. In this figure we assume that $\vdash^{dsk}$ is used and that the bidirectional system uses the GEN2* rule. Solid lines correspond to unrestricted relations between type systems, shadowed lines correspond to relations where the terms are unannotated and the contexts contain only Damas-Milner types.

4 A formalised type inference algorithm

In this section, we give a precise but abstract specification of a type inference *algorithm*. The final version of the bidirectional system was a syntax-directed system. A syntax-directed system is an important step towards a type inference algorithm because the steps of the algorithm could be driven by the syntax of the term, rather than having to search for a valid derivation.

However, a syntax-directed type system does not fully specify an inference algorithm. At certain points in the syntax-directed system, guessing is still required—for example, in the rule INST, the rules do not specify what types $\overline{\tau}$ should be used to instantiate the bound variables of a polytype. Because of this guess, typing is non-deterministic. By making different choices for $\overline{\tau}$ we can show that a given term has many different types.

The point of a type inference algorithm is to choose, out of all of these possible types, the one that is the "best" or most-general. Below, we formally specify a type inference algorithm for the bidirectional Odersky/Läufer system, based on the Damas-Milner "Algorithm W". We begin by discussing type variables (Section 4.1) and unification (Section 4.2). Then we give the formalisation of Algorithm W in Section 4.3 and finally extend it to higher-rank types in Section 4.4.

4.1 Type variables and substitutions

In the discussion so far we have encountered two distinct kinds of type variables: *ordinary* type variables and *meta* type variables. Consider the syntax of Damas-Milner types:

$$\begin{array}{rcl} \sigma & ::= & \forall\overline{a}.\tau \\ \tau & ::= & \texttt{Int} \mid \tau_1 \rightarrow \tau_2 \mid a \end{array}$$

The type variable "a" is part of the concrete syntax of types: $a \rightarrow \texttt{Int}$ and $\forall a.a \rightarrow a$ are both legal types. On the other hand, "τ" and "σ" are *meta-variables*, part of the language that we use to discuss types, but not part of the language of syntax of types themselves. For example, $\tau \rightarrow \tau$ is not itself a legal type. The typing judgements for a type system (Figure 13, for example) uses both kinds of variables. It uses "a" to mean "a type variable", and "τ" to mean "some type obeying the syntax of τ-types".

In a type inference algorithm, however, meta type variables are represented explicitly. The Algorithm W approach works as follows:

- When we must "guess" a monotype, such as in rule INST, we make up a fresh meta type variable, α.

- We carry around an idempotent substitution that maps meta type variables to monotypes (possibly involving other meta type variables).

- As the algorithm progresses, we generate equality constraints, which we solve by unification, extending the current substitution to reflect this solution.

For example, consider the application `reverse [1,2]`, where $\texttt{reverse} :: \forall a.[a] \rightarrow [a]$. We can infer the type of the application as follows. First, we instantiate the type of `reverse` with a fresh meta type variable, say β, yielding the type $[\beta] \rightarrow [\beta]$. Now, infer the type of `[1,2]`, yielding $[\texttt{Int}]$. Now, since `reverse` is applied to that list, we know that the equation $[\beta] = [\texttt{Int}]$ must hold. We can solve this equation by the standard unification algorithm, yielding the substitution $[\beta \mapsto \texttt{Int}]$.

To summarise, the basic infrastructure required by this approach is as follows:

- We distinguish between ordinary type variables (written a, b, c), and meta type variables (written α, β, γ).

- We need a source of fresh meta and ordinary type variables. The reason that we also require fresh ordinary type variables is that whenever we are going "inside" a polytype we need to treat the bound variables in the body of the type as completely fresh; therefore we need to replace them with fresh type variables. This point will become more clear in Section 4.3.

- We thread an ever-growing, idempotent substitution through the algorithm. This substitution is a finite map, that maps meta type variables (only!) to monotypes.

- We need a unification algorithm that takes the current substitution, and an equation between monotypes, and extends the substitution to make the two types equal. Indeed, we use the term "unifier" and "substitution" interchangeably.

The fact that meta type variables range only over monotypes is because our system is predicative: in rule INST for example we only "guess" a τ-type, not a σ-type. The syntax $ftv(\cdot)$ still denotes all free variables in the argument—meta and ordinary. Sometimes we use $fmv(\cdot)$ to denote the free meta type variables of the argument and $fov(\cdot)$ to denote the set of ordinary type variables of the argument.

4.2 Unification

In Figure 18 we give a unification procedure. It is written using inference rules, but it can be read very directly as an algorithm. We present it here primarily to introduce the notation; the algorithm itself is completely standard.

We give first-order unification in Figure 18. The inference rules can be seen as a procedure that, given an initial unifier S_0 and two types τ_1 and τ_2 returns a new substitution S_1—which extends S_0—and unifies the two types.

4.3 Algorithmic version of Damas-Milner type inference

Before doing type inference for higher-rank types, we begin by treating the original Damas-Milner system. Figures 16 and 17 show the type inference algorithm for Damas-Milner type inference. These rules are closely based on Figure 14: each rule in that figure has a corresponding rule in the algorithmic version.

The main judgement of the algorithm has the form

$$(S_0, \mathcal{A}_0) ; \Gamma \vdash t : \rho \Rightarrow (S_1, \mathcal{A}_1)$$

meaning that "given context Γ, an initial substitution S_0, a symbol supply $\mathcal{A}_0$ and term t, the algorithm produces the type ρ, substitution S_1 and a set of remaining symbols $\mathcal{A}_1$. This judgement relies on auxillary judgements for generalisation, instantiation and subsumption that correspond to the other judgements of Figure 14.

The symbol supply $\mathcal{A}$ is an unordered collection of distinct type variables, and models the supply of "fresh" type variables that is required by Algorithm W. Each judgement that needs fresh type variables takes a symbol supply $\mathcal{A}_0$ as input, and produces a depleted supply $\mathcal{A}_1$ as output. The notation $\mathcal{A}\mathcal{X}$ is the disjoint union of a finite set $\mathcal{X}$ and a supply $\mathcal{A}$.

In a similar way, most judgements take as input a substitution S_0 and return an extended substitution S_1. Unlike some presentations, we do not require that the returned type is a fixed point of the returned substitution.

The growing substitution and diminishing symbol supply are "threaded" through each judgement. For example, in rule APP of Figure 16, the incoming $(S_0, \mathcal{A}_0)$ are used to infer the type of the function t; that

$$\boxed{(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash t : \rho \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\rho \text{ output})$$

$$\frac{}{(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash i : \mathtt{Int} \Rightarrow (S_0, \mathcal{A}_0)} \text{ INT}$$

$$\frac{x : \sigma \in \Gamma \qquad (S_0, \mathcal{A}_0) \vdash^{inst} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash x : \rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ VAR}$$

$$\frac{(S_0, \mathcal{A}_0) \,;\, \Gamma, x : \beta \vdash t : \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0 \beta) \,;\, \Gamma \vdash \backslash x . t : \beta \to \rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ ABS}$$

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \,;\, \Gamma \vdash t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \,;\, \Gamma \vdash u : \rho_2 \Rightarrow (S_2, \mathcal{A}_2) \\ S_2 \vdash \rho_1 = \rho_2 \to \alpha \Rightarrow S_3 \end{array}}{(S_0, \mathcal{A}_0 \alpha) \,;\, \Gamma \vdash t \; u : \alpha \Rightarrow (S_3, \mathcal{A}_2)} \text{ APP}$$

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \,;\, \Gamma \vdash^{poly} u : \sigma \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \,;\, \Gamma, x : \sigma \vdash t : \rho \Rightarrow (S_2, \mathcal{A}_2) \end{array}}{(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash \mathtt{let} \; x \; \mathtt{=} \; u \; \mathtt{in} \; t : \rho \Rightarrow (S_2, \mathcal{A}_2)} \text{ LET}$$

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \,;\, \Gamma \vdash^{poly} t : \sigma' \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \vdash^{sh} \sigma' \leq \sigma \Rightarrow (S_2, \mathcal{A}_2) \\ (S_2, \mathcal{A}_2) \vdash^{inst} \sigma \leq \rho \Rightarrow (S_3, \mathcal{A}_3) \end{array}}{(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash (t :: \sigma) : \rho \Rightarrow (S_3, \mathcal{A}_3)} \text{ ANNOT}$$

Figure 16: Type inference algorithm for Damas-Milner system

returns $(S_1, \mathcal{A}_1)$ which are used in inferring the type of the argument u; and the result $(S_2, \mathcal{A}_2)$ is returned from the application.

Figure 17 gives the judgements for instantiation, generalisation, and subsumption (compare to Figure 14). The generalisation inference judgement

$$(S_0, \mathcal{A}_0) \,;\, \Gamma \vdash^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$$

returns an inferred polytype σ. Note that in the generalisation rule GEN we gather all the free *meta* type variables of the returned type that are not in the context — hence the $\overline{\alpha}$ notation there. When gathering these free meta type variables, we must be careful to appply the latest substitution S_1, which reflects all the constraints we have encountered so far. Any meta type variables $\overline{\alpha}$ that are mentioned *only* in $S_1(\rho)$ can be generalised, by replacing them with ordinary type variables $\overline{b}$, and then quantifying over $\overline{b}$. We do not extend S_1 with this latter substitution; instead, we simply substitute in $S_1(\rho)$.

It is worth discussing the rule SKOL a little more.

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \vdash^{sh} \sigma \leq [\overline{a \mapsto b}]\rho \Rightarrow (S_1, \mathcal{A}_1) \\ \overline{b} \notin ftv(S_1 \sigma) \qquad \overline{b} \notin ftv(S_1(\forall \overline{a} . \rho)) \end{array}}{(S_0, \mathcal{A}_0 \overline{b}) \vdash^{sub\sigma} \sigma \leq \forall \overline{a} . \rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ SKOL}$$

$$\boxed{(S_0, \mathcal{A}_0)\,;\Gamma \vdash^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\sigma \text{ output})$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0)\,;\Gamma \vdash t : \rho \Rightarrow (S_1, \mathcal{A}_1)\\ \overline{\alpha} = ftv(S_1\rho) - ftv(S_1\Gamma)\end{array}}{(S_0, \mathcal{A}_0\overline{b})\,;\Gamma \vdash^{poly} t : \forall \overline{b}.[\overline{\alpha} \mapsto \overline{b}]S_1\rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ GEN}$$

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{inst} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\rho \text{ output})$$

$$\frac{}{(S_0, \mathcal{A}_0\overline{\beta}) \vdash^{inst} \forall \overline{a}.\rho \leq [\overline{a \mapsto \beta}]\rho \Rightarrow (S_0, \mathcal{A}_0)} \text{ INST}$$

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{sh} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\sigma_1, \sigma_2 \text{ inputs})$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0) \vdash^{sh} \sigma \leq [\overline{a \mapsto b}]\rho \Rightarrow (S_1, \mathcal{A}_1)\\ \overline{b} \notin ftv(S_1\sigma) \quad \overline{b} \notin ftv(S_1(\forall \overline{a}.\rho))\end{array}}{(S_0, \mathcal{A}_0\overline{b}) \vdash^{sh} \sigma \leq \forall \overline{a}.\rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ SKOL}$$

$$\frac{(S_0, \mathcal{A}_0) \vdash^{sh} [\overline{a \mapsto \beta}]\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0\overline{\beta}) \vdash^{sh} \forall \overline{a}.\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)} \text{ SPEC}$$

$$\frac{S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1}{(S_0, \mathcal{A}_0) \vdash^{sh} \tau_1 \leq \tau_2 \Rightarrow (S_1, \mathcal{A}_0)} \text{ MONO}$$

Figure 17: Generalisation, instantiation and subsumption for Damas-Milner

For this rule we need to check if σ is more polymorphic than $\forall \overline{a}.\rho$. The first thing that we need to confirm is that σ is more polymorphic than $[\overline{a \mapsto b}]\rho$, that is, the body of $\forall \overline{a}.\rho$ where we have replaced the bound variables with completely fresh type variables—hence the requirement for the symbol supply to also contain fresh ordinary type variables. Intuitively this step assumes any unknown instantiation of $\forall \overline{a}.\rho$ and tries to ensure that σ is more polymorphic that this instatiation. The algorithm will yield back a unifier S_1, that may possibly contain $\overline{b}$ in its range—consider for example the returned unifier that takes some of the meta variables for an instantiation of σ to monotypes containing $\overline{b}$. However none of the **free meta type variables** of $S_1\sigma$ and $S_1(\forall \overline{a}.\rho)$ should be among $\overline{b}$, because then these variables would escape their scope: We do not want to allow unbound ordinary type variables in our returned types.

These rules form an algorithm because there is no guessing to be done. Not only is the derivation constrained by the syntax of the term, but all guessing has been eliminated. For example, the guessing of the argument type in the rule ABS is replaced with the generation of a fresh meta variable. Likewise, the guessing of $\overline{\tau}$ in the rules INST and SPEC uses a list of fresh meta variables instead. When types must be compared for equality (in the rules APP and MONO) the algorithm instead uses unification to determine if there is some substition for the metavariables that makes these types equal.

$$\boxed{S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1}$$

$$(\tau_1, \tau_2 \text{ inputs})$$

$$\frac{\tau = \alpha \text{ or } \tau = a}{S_0 \vdash \tau = \tau \Rightarrow S_0} \text{ UREFL}$$

$$\frac{\begin{array}{c} \alpha \in dom(S_0) \\ S_0 \vdash S_0\alpha = \tau \Rightarrow S_1 \end{array}}{S_0 \vdash \alpha = \tau \Rightarrow S_1} \text{ BVAR1} \qquad \frac{\begin{array}{c} \alpha \in dom(S_0) \\ S_0 \vdash S_0\alpha = \tau \Rightarrow S_1 \end{array}}{S_0 \vdash \tau = \alpha \Rightarrow S_1} \text{ BVAR2}$$

$$\frac{\begin{array}{c} \alpha \notin dom(S_0) \\ \alpha \notin ftv(S_0\tau) \end{array}}{S_0 \vdash \alpha = \tau \Rightarrow [\alpha \mapsto S_0\tau] \cdot S_0} \text{ UVAR1} \qquad \frac{\begin{array}{c} \alpha \notin dom(S_0) \\ \alpha \notin ftv(S_0\tau) \end{array}}{S_0 \vdash \tau = \alpha \Rightarrow [\alpha \mapsto S_0\tau] \cdot S_0} \text{ UVAR2}$$

$$\frac{S_0 \vdash \tau_1 = \tau_1' \Rightarrow S_1 \quad S_1 \vdash \tau_2 = \tau_2' \Rightarrow S_2}{S_0 \vdash \tau_1 \rightarrow \tau_2 = \tau_1' \rightarrow \tau_2' \Rightarrow S_2} \text{ UFUN}$$

Figure 18: Unification

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{\rightarrow} \rho = \sigma \rightarrow \sigma' \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\rho \text{ input}, \sigma \rightarrow \sigma' \text{ output})$$

$$\frac{S_0 \vdash \gamma = \alpha \rightarrow \beta \Rightarrow S_1}{(S_0, \mathcal{A}\alpha\beta) \vdash^{\rightarrow} \gamma = \alpha \rightarrow \beta \Rightarrow (S_1, \mathcal{A})} \text{ UFUN-MONO}$$

$$\frac{}{(S_0, \mathcal{A}_0) \vdash \sigma \rightarrow \sigma' = \sigma \rightarrow \sigma' \Rightarrow (S_0, \mathcal{A}_0)} \text{ UFUN-ARROW}$$

Figure 19: Arrow Unification

4.4 Algorithmic version of the bidirectional system

We now extend algorithm W to the bidirectional type system. The revised algorithm appears in Figures 21-24.

The first important difference between the previous algorithm and this one is the definition of subsumption. The new definition follows closely canonical derivations of $\vdash^{dsk}$. The new subsumption relation is given in Figure 22.

In addition to the unification procedure we gave in previous section we also require a procedure that unifies a type with an arrow type. The arrow unification judgement of the form

$$(S_0, \mathcal{A}_0) \vdash^{\rightarrow} \rho = \sigma_1 \rightarrow \rho_1 \Rightarrow (S_1, \mathcal{A}_1)$$

takes an initial unifier S_0, an initial supply $\mathcal{A}_0$ and a type ρ. It produces a bigger unifier S_1 and an arrow type that matches ρ once S_1 has been applied to these types. The rules are straightforward and given in Figure 19.

$$\boxed{\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \sigma' \Rightarrow \mathcal{A}_1}$$
$$(\sigma' \text{ output})$$

$$\frac{\tau = \alpha \text{ or } \tau = a}{\mathcal{A}_0 \vdash^{pr} \tau \mapsto \tau \Rightarrow \mathcal{A}_0} \text{ PRMONO}$$

$$\frac{\mathcal{A}_0 \vdash^{pr} \sigma_2 \mapsto \forall \overline{a}.\rho_2 \Rightarrow \mathcal{A}_1}{\mathcal{A}_0 \vdash^{pr} \sigma_1 \rightarrow \sigma_2 \mapsto \forall \overline{a}.\sigma_1 \rightarrow \rho_2 \Rightarrow \mathcal{A}_1} \text{ PRFUN}$$

$$\frac{\mathcal{A}_0 \vdash^{pr} \overline{[a \mapsto b]}\rho \mapsto \forall \overline{c}.\rho' \Rightarrow \mathcal{A}_1}{\mathcal{A}_0 \overline{b} \vdash^{pr} \forall \overline{a}.\rho \mapsto \forall \overline{b}\overline{c}.\rho' \Rightarrow \mathcal{A}_1} \text{ PRPOLY}$$

Figure 20: Weak prenex conversion

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{inst}_\delta \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)}$$
$$(\rho \text{ output when } \delta = \Uparrow, \text{ input when } \delta = \Downarrow)$$

$$\frac{}{(S_0, \mathcal{A}_0 \overline{\beta}) \vdash^{inst}_\Uparrow \forall \overline{a}.\rho \leq \overline{[a \mapsto \beta]}\rho \Rightarrow (S_0, \mathcal{A}_0)} \text{ AINST1}$$

$$\frac{(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0) \vdash^{inst}_\Downarrow \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)} \text{ AINST2}$$

Figure 21: Algorithmic instantiation

Weak-prenex conversion is given by the following judgement:

$$\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \sigma' \Rightarrow \mathcal{A}_1$$

The rules are given in Figure 20; notice that it consumes symbols from the supply in order to "freshen" the quantified variables of the types.

The instantiation relation has an algorithmic version, shown in Figure 21. The instantiation judgements are given by $(S_0, \mathcal{A}_0) \vdash^{inst}_\delta \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)$. Again, they take an initial unifier S_0 and a supply $\mathcal{A}_0$ and either check or infer that σ is more polymorphic than ρ and return the rest of the supply $\mathcal{A}_1$, extending perhaps the unifier to S_1.

The inference and checking judgements are given by $(S_0, \mathcal{A}_0); \Gamma \vdash_\delta t : \rho \Rightarrow (S_1, \mathcal{A}_1)$. In the case of inference, just like in the Damas-Milner algorithm, we take a unifier S_0 and a supply $\mathcal{A}_0$ and, in return back a type ρ. In checking mode we check if we can assign the ρ type to the term t. In any case we return an extended unifier, S_1, as well as the rest of the supply, $\mathcal{A}_1$. Notice the use of arrow unification in the abstraction checking judgements ALG-ABS1 and ALG-ABS2 as well as the application judgement ALG-APP.

Finally we have the polytype inference and checking judgements $(S_0, \mathcal{A}_0); \Gamma \vdash^{poly}_\delta t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$. The inference case is very alike the generalisation inference case of the Damas-Milner algorithm. The corresponding checking judgement should be read as: "check that σ is at most as polymorphic as the term t". Notice the similarity of this rule compared to the SKOL rule of the subsumption judgement. We also have

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)}$$

$$(\sigma_1, \sigma_2 \text{ inputs})$$

$$\frac{\begin{array}{c} \mathcal{A}_0 \vdash^{pr} \sigma_2 \mapsto \forall \overline{a} . \rho \Rightarrow \mathcal{A}_1 \\ (S_0, \mathcal{A}_1) \vdash^{dsk*} \sigma_1 \leq \rho \Rightarrow (S_1, \mathcal{A}_2) \\ \overline{a} \notin ftv(S_1\sigma_1) \quad \overline{a} \notin ftv(S_1\sigma_2) \end{array}}{(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_2)} \text{ ASKOL}$$

$$\boxed{(S_0, \mathcal{A}_0) \vdash^{dsk*} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)}$$

$$\frac{(S_0, \mathcal{A}_0) \vdash^{dsk*} \overline{[a \mapsto \beta]}\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0 \overline{\beta}) \vdash^{dsk*} \forall \overline{a} . \rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)} \text{ ASPEC}$$

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \vdash^{\rightarrow} \rho = \sigma_1 \rightarrow \sigma_2 \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_3 \leq \sigma_1 \Rightarrow (S_2, \mathcal{A}_2) \\ (S_2, \mathcal{A}_2) \vdash^{dsk*} \sigma_2 \leq \sigma_4 \Rightarrow (S_3, \mathcal{A}_3) \end{array}}{(S_0, \mathcal{A}_0) \vdash^{dsk*} \rho \leq \sigma_3 \rightarrow \sigma_4 \Rightarrow (S_3, \mathcal{A}_3)} \text{ AFUN1}$$

$$\frac{\begin{array}{c} (S_0, \mathcal{A}_0) \vdash^{\rightarrow} \rho = \sigma_3 \rightarrow \sigma_4 \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_3 \leq \sigma_1 \Rightarrow (S_2, \mathcal{A}_2) \\ (S_2, \mathcal{A}_2) \vdash^{dsk*} \sigma_2 \leq \sigma_4 \Rightarrow (S_3, \mathcal{A}_3) \end{array}}{(S_0, \mathcal{A}_0) \vdash^{dsk*} \sigma_1 \rightarrow \sigma_2 \leq \rho \Rightarrow (S_3, \mathcal{A}_3)} \text{ AFUN2}$$

$$\frac{S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1}{(S_0, \mathcal{A}_0) \vdash^{dsk*} \tau_1 \leq \tau_2 \Rightarrow (S_1, \mathcal{A}_0)} \text{ AMONO}$$

Figure 22: Algorithmic subsumption

to peform weak-prenex conversion and make sure that the returned unifier does not unify some of the free meta variables of the context or the polytype with these fresh variables.

4.5 Properties of the type inference algorithm

The *soundness* theorem states that if the algorithm determines that a term t has a type ρ, then there is a derivation using the rules in Figure 9 assigning the same type to t.
Theorem 4.1 (Soundness).

1. *Suppose $\mathcal{A}_0 \notin ftv(\Gamma)$ and $([], \mathcal{A}_0) ; \Gamma \vdash_{\Uparrow} t : \rho \Rightarrow (S, \mathcal{A}_1)$. Then for any substitution V with $dom(V) = fmv(S\Gamma, S\rho)$ we have $VS\Gamma \vdash_{\Uparrow} t : VS\rho$.*

2. *Suppose $\mathcal{A}_0 \notin ftv(\Gamma, \rho)$ and $([], \mathcal{A}_0) ; \Gamma \vdash_{\Downarrow} t : \rho \Rightarrow (S, \mathcal{A}_1)$. Then for any substitution V with $dom(V) = fmv(S\Gamma, S\rho)$ we have $VS\Gamma \vdash_{\Downarrow} t : VS\rho$.*

The role of the substitution V is auxiliary and at first reading one can completely ignore it. The reason is that the algorithm may return some un-unified meta type variables. For example $([], \beta) ; \vdash \backslash x . x : \beta \rightarrow \beta \Rightarrow ([], \{\})$.

$$\boxed{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\delta t : \rho \Rightarrow (S_1, \mathcal{A}_1)}$$

(ρ output when $\delta = \Uparrow$, input when $\delta = \Downarrow$)

$$\frac{x : \sigma \in \Gamma \quad (S_0, \mathcal{A}_0) \vdash_\delta^{inst} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\delta x : \rho \Rightarrow (S_1, \mathcal{A}_1)} \; \text{ALG-VAR}$$

$$\frac{(S_0, \mathcal{A}_0)\,;\Gamma, x : \beta \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0\beta)\,;\Gamma \vdash_\Uparrow \backslash x\,.\,t : \beta \rightarrow \rho \Rightarrow (S_1, \mathcal{A}_1)} \; \text{ALG-ABS1}$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0) \vdash^\rightarrow \rho = \sigma_a \rightarrow \sigma_r \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1)\,;\Gamma, x : \sigma_a \vdash_\Downarrow^{poly} t : \sigma_r \Rightarrow (S_2, \mathcal{A}_2)\end{array}}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Downarrow \backslash x\,.\,t : \rho \Rightarrow (S_2, \mathcal{A}_2)} \; \text{ALG-ABS2}$$

$$\frac{(S_0, \mathcal{A}_0)\,;\Gamma, x : \sigma \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow \backslash x :: \sigma\,.\,t : \sigma \rightarrow \rho \Rightarrow (S_1, \mathcal{A}_1)} \; \text{ALG-AABS1}$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0) \vdash^\rightarrow \rho = \sigma_a \rightarrow \sigma_r \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_a \leq \sigma_x \Rightarrow (S_2, \mathcal{A}_2) \\ (S_2, \mathcal{A}_2)\,;\Gamma, x : \sigma_x \vdash_\Downarrow^{poly} t : \sigma_r \Rightarrow (S_3, \mathcal{A}_3)\end{array}}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Downarrow \backslash x :: \sigma_x\,.\,t : \rho \Rightarrow (S_3, \mathcal{A}_3)} \; \text{ALG-AABS2}$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1) \vdash^\rightarrow \rho_1 = \sigma \rightarrow \sigma' \Rightarrow (S_2, \mathcal{A}_2) \\ (S_2, \mathcal{A}_2)\,;\Gamma \vdash_\Downarrow^{poly} u : \sigma \Rightarrow (S_3, \mathcal{A}_3) \\ (S_3, \mathcal{A}_3) \vdash_\delta^{inst} \sigma' \leq \rho \Rightarrow (S_4, \mathcal{A}_4)\end{array}}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\delta t\,u : \rho \Rightarrow (S_4, \mathcal{A}_4)} \; \text{ALG-APP}$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow^{poly} u : \sigma \Rightarrow (S_1, \mathcal{A}_1) \\ (S_1, \mathcal{A}_1)\,;\Gamma, x : \sigma \vdash_\delta t : \rho \Rightarrow (S_2, \mathcal{A}_2)\end{array}}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\delta \texttt{let } x \texttt{ = } u \texttt{ in } t : \rho \Rightarrow (S_2, \mathcal{A}_2)} \; \text{ALG-LET}$$

Figure 23: Inference/Checking Algorithm

Of course meta type variables do not appear syntactically in the bidirectional system. The substitution V simply eliminates such meta variables from the returned types.

Completeness on the other hand says that if the bidirectional system assigns the type ρ to the term t, then the algorithm can infer a type ρ' that can produce ρ through some substitution.

Again, when we state this theorem, we must constrain the the symbol generator to be "fresh" from the variables in the judgement.

Theorem 4.2 (Completeness). *The algorithm is complete with respect to the syntax-directed system.*

1. *Suppose that $\mathcal{A}_0 \notin ftv(\Gamma, \rho)$. If $\Gamma \vdash_\Uparrow t : \rho$ then $([], \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow t : \rho' \Rightarrow (S, \mathcal{A}_1)$ and there exists R such that $RS\rho' = \rho$.*

2. *Suppose that $\mathcal{A}_0 \notin ftv(\Gamma, \rho)$. If $\Gamma \vdash_\Downarrow t : \rho$ then $([], \mathcal{A}_0)\,;\Gamma \vdash_\Downarrow t : \rho \Rightarrow (S, \mathcal{A}_1)$ and $S\rho = \rho$, $S\Gamma = \Gamma$.*

$$\boxed{(S_0, \mathcal{A}_0)\,;\Gamma \vdash^{poly}_\delta\ t : \sigma \Rightarrow (S_1, \mathcal{A}_1)}$$
$$(\sigma \text{ output when } \delta = \Uparrow, \text{ input when } \delta = \Downarrow)$$

$$\frac{\begin{array}{c}(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow\ t : \rho \Rightarrow (S_1, \mathcal{A}_1) \\ \overline{\alpha} = \mathit{ftv}(S_1\rho) - \mathit{ftv}(S_1\Gamma)\end{array}}{(S_0, \mathcal{A}_0\overline{b})\,;\Gamma \vdash^{poly}_\Uparrow\ t : \forall \overline{b}\,.\,[\overline{\alpha} \mapsto \overline{b}]S_1\rho \Rightarrow (S_1, \mathcal{A}_1)}\ \text{ALG-GEN1}$$

$$\frac{\begin{array}{c}\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall \overline{a}\,.\,\rho \Rightarrow \mathcal{A}_1 \\ (S_0, \mathcal{A}_1)\,;\Gamma \vdash_\Downarrow\ t : \rho \Rightarrow (S_1, \mathcal{A}_2) \\ \overline{a} \notin \mathit{ftv}(S_1\Gamma) \quad \overline{a} \notin \mathit{ftv}(S_1\sigma)\end{array}}{(S_0, \mathcal{A}_0)\,;\Gamma \vdash^{poly}_\Downarrow\ t : \sigma \Rightarrow (S_1, \mathcal{A}_2)}\ \text{ALG-GEN2}$$

Figure 24: Algorithmic generalisation

Combining completeness and soundness gives us a principal types property for the bidirectional system (see also Section 4.7). It tells us that out of all the types that bidirectional system assigns to a term, there is a best one (the principal one), such that all others are substitution instances of that one. That type is precisely the one picked out by the algorithm.

4.6 Proofs about the algorithm

Definition 4.3 (Excluded-$\mathcal{X}$ substitution equivalence). *Given a set of variables $\mathcal{X}$, we define the* excluded-$\mathcal{X}$ equivalence *relation on substitutions as:*

$$S_1 = S_2\backslash_\mathcal{X} \quad \Leftrightarrow \quad \forall a \notin \mathcal{X}, S_1(a) = S_2(a)$$

Intuitively, two substitutions are excluded-$\mathcal{X}$ equivalent if they agree everywhere except perhaps for some variables in $\mathcal{X}$. Recall that we write $\mathcal{X}_1 \notin \mathcal{X}_2$, where $\mathcal{X}_1, \mathcal{X}_2$ are sets of variables, meaning that the two sets are disjoint. A *unifier* is a substitution whose domain contains only **meta type variables**. In what follows symbols S, T, U, V denote unifiers unless stated explicitly otherwise. When we write $[a \mapsto b] \cdot S$, since a can't be in the domain of a unifier S we mean the (renaming) substitutions of a for b in the range of S. On the other hand $[\alpha \mapsto b] \cdot S$ denotes the substitution composition of $[\alpha \mapsto b]$ and S.

Let us start by proving some sanity checks and useful facts about unification. The first property that we need is that when we start with a well-defined substitution, we end up with a well-defined substitution: For our purposes the notion of a mathematically well-defined substitution will coincide with idempotency:

$$S \text{ is well-defined iff } \forall \sigma\,.\,S\sigma = S(S\sigma)$$

Lemma 4.4 (Idempotency of unifiers). *If S_0 is idempotent and $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ then S_1 is idempotent.*

Proof. The proof is by induction on the definition of the unification procedure. The case for UREFL is trivial. The cases for BVAR1, BVAR2, UFUN follow easily by applying the induction hypotheses. Let us consider the case UVAR1 (the case UVAR2 is similar). Here we have that $S_0 \vdash \alpha = \tau \Rightarrow S_1$, given that

$$\alpha \notin \mathit{dom}(S_0) \tag{1}$$
$$\alpha \notin \mathit{ftv}(S_0\tau) \tag{2}$$
$$S_1 = [\alpha \mapsto S_0\tau] \cdot S_0 \tag{3}$$

We show by induction on σ that $S_1 S_1 \sigma = S_1 \sigma$. All the cases are easy or follow by induction hypothesis except for the case where $\sigma = \beta$. If $\beta = \alpha$ we have $S_1 \alpha = S_0 \tau$, which means that:

$$
\begin{aligned}
S_1 S_1 \alpha &= S_1 S_0 \tau \\
&= [\alpha \mapsto S_0 \tau] S_0 S_0 \tau \\
&= [\alpha \mapsto S_0 \tau] S_0 \tau \qquad \text{(because } S_0 \text{ is idempotent)} \\
&= S_0 \tau \qquad\qquad\quad \text{(by (2))} \\
&= S_1 \alpha
\end{aligned}
$$

Similarly, if $\beta \neq \alpha$ we have that

$$
\begin{aligned}
S_1 S_1 b &= [\alpha \mapsto S_0 \tau] S_0 [a \mapsto S_0 \tau] S_0 \beta \\
&= [\alpha \mapsto S_0 \tau] S_0 S_0 \beta \\
&= [\alpha \mapsto S_0 \tau] S_0 \beta \\
&= S_1 \beta
\end{aligned}
$$

$\square$

It is a series of easy inductions to show that all unifiers mentioned throughout the paper are idempotent. These proofs rely on the last lemma. We omit these proofs and assume that we only deal with well-defined unifiers in the rest of the document.

Now a quick check about the variables of the unifiers. What this says is that our returned unifier is larger than the input unifier and that it doesn't contain symbols made out of thin air.

Lemma 4.5. *If $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ then:*

1. $vars(S_1) \subseteq vars(S_0) \cup ftv(\tau_1, \tau_2)$

2. $dom(S_0) \subseteq dom(S_1)$

3. $range(S_1) \subseteq range(S_0) \cup ftv(\tau_1, \tau_2)$

Proof. Easy induction. $\square$

The next lemma establishes the soundness of unification.

Lemma 4.6 (Unification soundnesss). *If $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ then $S_1 \tau_1 = S_1 \tau_2$ and there exists a substitution R such that $S_1 = R \cdot S$. Moreover $vars(R) \subseteq ftv(S_0 \tau_1, S_0 \tau_2)$.*

Proof. The proof is by induction on the derivation of unification. We examine all cases.

- Case UREFL. Trivially take R to be the empty substitution.

- Case UFUN. Here we have that $S_0 \vdash \tau_1 \to \tau_2 = \tau_1' \to \tau_2' \Rightarrow S_2$ given that $S_0 \vdash \tau_1 = \tau_1' \Rightarrow S_1$ and $S_1 \vdash \tau_1' = \tau_2' \Rightarrow S_2$. By induction hypothesis $S_1 \tau_1 = S_1 \tau_1'$ and $S_1 = R_1 \cdot S_0$ for some R_1 and $vars(R_1) \subseteq ftv(S_0 \tau_1, S_0 \tau_1')$. Also by induction hypothesis $S_2 \tau_2 = S_2 \tau_2'$ and $S_2 = R_2 \cdot S_1$ for some R_2. Then we have that

$$
\begin{aligned}
& vars(R_2) \subseteq ftv(S_1 \tau_2, S_1 \tau_2') \\
\Rightarrow \quad & vars(R_2) \subseteq R_1(ftv(S_0 \tau_2, S_0 \tau_2')) \\
\Rightarrow \quad & vars(R_2) \subseteq ftv(S_0 \tau_1, S_0 \tau_1') \cup ftv(S_0 \tau_2, S_0 \tau_2')
\end{aligned}
$$

  Then we have $R_2 S_1 \tau_1 = R_2 S_1 \tau_1'$, hence $S_2 \tau_1 = S_2 \tau_1'$ and taking $R = R_2 \cdot R_1$ finishes the case since $vars(R) \subseteq vars(R_1, R_2)$ as well.

- Case BVAR1 (the case for BVAR2 is similar). Here we have $S_0 \vdash \alpha = \tau \Rightarrow S_1$ given that $S_0 \vdash S_0\alpha = \tau \Rightarrow S_1$ when $\alpha \in dom(S_0)$. By induction hypothesis $S_1 S_0 \alpha = S_1 \tau$ and $S_1 = R \cdot S_0$ for some R. Moreover $vars(R) \subseteq ftv(S_0 S_0\alpha, S_0\tau)$ and because of idempotency $vars(R) \subseteq ftv(S_0\alpha, S_0\tau)$. Then $RS_0 S_0 \alpha = S_1 \tau$ and because S_0 is idempotent $RS_0\alpha = S_1\tau$ or $S_1\alpha = S_1\tau$. Therefore taking the same R finishes the case.

- Case UVAR1 (the case for UVAR2 is similar). Here $S_0 \vdash \alpha = \tau \Rightarrow [\alpha \mapsto S_0\tau] \cdot S_0$ given that

$$\alpha \notin dom(S_0) \tag{1}$$
$$\alpha \notin ftv(S_0\tau) \tag{2}$$

Then $[\alpha \mapsto S_0\tau]S_0\alpha = S_0\tau$ because of (1) and $S_0\tau = [\alpha \mapsto S_0\tau]S_0\tau$ because of (2). Finally pick $R = [\alpha \mapsto S_0\tau]$. Then $vars(R) = ftv(\alpha, S_0\tau) = ftv(S_0\alpha, S_0\tau)$.

$\square$

Lemma 4.7. *If $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ then $fov(S_1\tau_1, S_1\tau_2) \subseteq fov(S_0\tau_1, S_0\tau_2)$.*

Proof. By unification soundness, Lemma 4.6, we have that $fov(S_1\tau_1, S_1\tau_2) = fov(S_1\tau_1) = fov(RS_0\tau_1)$ and we know that $fov(RS_0\tau_1) \subseteq fov(R) \cup fov(S_0\tau_1) \subseteq fov(S_0\tau_1, S_0\tau_2)$. $\square$

Next we establish completeness of unification, that is, we will show that if two types are unified by some substitution, then our algorithm returns always a most general unifier.

Lemma 4.8 (Unification completeness). *If $SS_0\tau_1 = SS_0\tau_2$ then unification of S_0, τ_1 and τ_2 succeeds, that is, $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ such that $S \cdot S_0 = R \cdot S_1$ for some R. Moreover $vars(S_1) \subseteq vars(S_0), ftv(\tau_1, \tau_2)$ and $vars(R) \subseteq vars(S, S_1, S_0)$.*

Proof. Consider the following lexicographic pair to be a metric for a given unifier S_0, types τ_1 and τ_2.

$$\mu = \langle \mid range(S_0) \cup ftv(\tau_1, \tau_2) \mid, size(\tau_1) + size(\tau_2)\rangle$$

We show unification completeness by induction on the value of μ (We can use the same metric to show termination of unification as well). Observe first of all that the two types cannot be tow distinct type variables. We proceed by case analysis on the possible forms of τ_1 and τ_2.

- Both of them are arrow types, that is $\tau_1 = \tau_{11} \rightarrow \tau_{12}$ and $\tau_2 = \tau_{21} \rightarrow \tau_{22}$. In this case we have that $SS_0\tau_{11} = SS_0\tau_{21}$ and $SS_0\tau_{12} = SS_0\tau_{22}$. By induction $S_0 \vdash \tau_{11} = \tau_{21} \Rightarrow S_1$ such that for some R_1 we have $S \cdot S_0 = R_1 \cdot S_1$. Now also by induction hypothesis (by Lemma 4.5 $range(S_1) \subseteq range(S_0) \cup ftv(\tau_{11}, \tau_{21})$) $S_1 \vdash \tau_{12} = \tau_{22} \Rightarrow S_2$ such that $R_1 \cdot S_1 = R \cdot S_2$ for some R hence giving us $S \cdot S_0 = R \cdot S_2$ hence giving us the result by UFUN.

- Assume that they both are ordinary variables: Then they have to be the same since the only rule applicable would be UREFL and the case follows trivially.

- One of them is a meta variable and assume $\tau_1 = \alpha$ (the case where τ_2 is a variable is symmetric). Assume also that $\tau_2 = \tau$. Then, if τ_2 is exactly the same variable then UREFL is applicable and the result follows easily.

  If $\tau_2 = \tau \neq \alpha$ we have the following cases:

  - $\alpha \in dom(S_0)$. In this case if we can show that $SS_0(S_0\alpha) = SS_0\tau$ then we will be done, as we will be able to apply the induction hypothesis ($\mid range(S_0) \cup ftv(\alpha, \tau) \mid \geq \mid range(S_0) \cup ftv(S_0\alpha, \tau) \mid = \mid range(S_0) \cup ftv(\tau) \mid$, since $\alpha \notin range(S_0)$ because S_0 is idempotent) to get that $S_0 \vdash S_0\alpha = \tau \Rightarrow S_1$ with $S \cdot S_0 = R \cdot S_1$ and by rule BVAR1 the case will be finished. But observe that the above follows directly from the idempotency of S_0.

– $\alpha \notin dom(S_0)$. Here we have the following cases for $S_0\tau$:

* $S_0\tau = \alpha$. Then since $\alpha \notin dom(S_0)$ we have that $\tau = \beta$ for some $\beta \in dom(S_0)$. But this means that we can apply rule BVAR2 with UREFL in its premise to get the result, taking $R = S$.

* $\alpha \in ftv(S_0\tau)$. This case can't happen as it would require $SS_0\alpha = SS_0\tau$ or $S\alpha = S(S_0\tau)$ but the sizes of the types in the left and right hand side can never be the same.

* $\alpha \notin ftv(S_0\tau)$. In this case we can apply the rule UVAR1 to get the result. Now we need to find R such that $S \cdot S_0 = R \cdot [\alpha \mapsto S_0\tau] \cdot S_0$. Since $SS_0\alpha = SS_0\tau$ it must be that $\alpha \in dom(S)$ and $S\alpha = S(S_0\tau)$, which means we can write S as $S = S' \cdot [\alpha \mapsto S_0\tau]$ since also $\alpha \notin ftv(S_0\tau)$. Now just pick $R = S'$ and we are done.

$\square$

From the soundness and completeness of the unification procedure we can derive two easy corrollaries for the arrow unification procedure.

Lemma 4.9 (Arrow unification soundness and completeness).

1. If for some $\mathcal{A}_0$ such that $\mathcal{A}_0 \notin ftv(\rho), vars(S_0)$ it is $(\mathcal{A}_0, S_0) \vdash \rho = \sigma \to \sigma' \Rightarrow (S_1, \mathcal{A}_1)$ then $S_1\rho = S_1\sigma \to S_1\sigma'$ and $\exists R$ such that $S_1 = R \cdot S_0$. Moreover $ftv(\sigma \to \sigma') \subseteq vars(S_0), ftv(\rho), (\mathcal{A}_0 - \mathcal{A}_1)$ and $vars(S_1) \subseteq vars(S_0), ftv(\rho), (\mathcal{A}_0 - \mathcal{A}_1)$ and $vars(R) \subseteq ftv(S_0\rho), (\mathcal{A}_0 - \mathcal{A}_1)$.

2. If $SS_0\rho = \sigma_0 \to \sigma_0'$ then for any symbol supply $\mathcal{A}_0$ such that $\mathcal{A}_0 \notin vars(S, S_0)$ and $\mathcal{A}_0 \notin ftv(\rho, \sigma_0, \sigma_0')$, $(\mathcal{A}_0, S_0) \vdash^{\to} \rho = \sigma \to \sigma' \Rightarrow (S_1, \mathcal{A}_1)$ such that $\exists R$ with $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$ and $RS_1\sigma = \sigma_0$ and $RS_1\sigma' = \sigma_0'$. Moreover $vars(R) \notin \mathcal{A}_1$ and $vars(S_1) \notin \mathcal{A}_1$.

Proof. The soundness part is by induction on the arrow unification definition and appealing to unification soundness, while the completeness part is by case analysis on the type structure and appealing to unification completeness. $\square$

Here's a lemma about the free ordinary variables during arrow unification.

Lemma 4.10. If $\mathcal{A}_0 \notin ftv(\rho), vars(S_0)$ and $(\mathcal{A}_0, S_0) \vdash^{\to} \rho = \sigma_1 \to \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ then $fov(S_1\rho, S_1\sigma_1, S_1\sigma_2) \subseteq fov(S_0\rho)$.

Proof. Like the proof of Lemma 4.7. $\square$

4.6.1 Completeness

In this section we show that the types that the type system attributes to terms can be considered "instances" of types that the algorithm discovers. We need a series of auxilliary lemmas first.

Lemma 4.11. If $\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall\overline{a}.\rho \Rightarrow \mathcal{A}_1$ then $\overline{a} \subseteq \mathcal{A}_0 - \mathcal{A}_1$, $ftv(\sigma) = ftv(\forall\overline{a}.\rho)$.

Proof. Easy induction on the definition of $\vdash^{pr}$. $\square$

Lemma 4.12.

1. If $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\sigma_1, \sigma_2)$ and $(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ then there exists R such that $S_1 = R \cdot S_0$, $vars(S_1) \subseteq vars(S_0) \cup ftv(\sigma_1, \sigma_2) \cup (\mathcal{A}_0 - \mathcal{A}_1)$, and $vars(R) \subseteq ftv(S_0\sigma_1, S_0\sigma_2) \cup (\mathcal{A}_0 - \mathcal{A}_1)$.

2. If $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\sigma_1, \sigma_2)$ and $(S_0, \mathcal{A}_0) \vdash^{dsk*} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ then there exists R such that $S_1 = R \cdot S_0$, $vars(S_1) \subseteq vars(S_0) \cup ftv(\sigma_1, \sigma_2) \cup (\mathcal{A}_0 - \mathcal{A}_1)$, and $vars(R) \subseteq ftv(S_0\sigma_1, S_0\sigma_2) \cup (\mathcal{A}_0 - \mathcal{A}_1)$.

Proof. The two claims can be proved simultaneously by induction on the height of the derivations, appealing to the unification properties in the monotype case. □

The next lemma asserts that no ordinary type variables from the symbol supply escape in the "useful" range of the unifiers.

Lemma 4.13.

1. If $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\sigma_1, \sigma_2)$, $(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$, and $fov(S_0\sigma_1, S_0\sigma_2) \subseteq \mathcal{X}$, then $fov(S_1\sigma_1, S_1\sigma_2) \subseteq \mathcal{X}$.

2. If $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\sigma_1, \sigma_2)$, $(S_0, \mathcal{A}_0) \vdash^{dsk*} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$, and $fov(S_0\sigma_1, S_0\sigma_2) \subseteq \mathcal{X}$, then $fov(S_1\sigma_1, S_1\sigma_2) \subseteq \mathcal{X}$.

Proof. We prove the two claims simultaneously by induction on height of the derivations.

For the first part we only have the case of ASKOL. We have $(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_2)$, given that

$$\mathcal{A}_0 \vdash^{pr} \sigma_2 \mapsto \forall \overline{a}.\rho \Rightarrow \mathcal{A}_1 \tag{1}$$

$$(S_0, \mathcal{A}_1) \vdash^{dsk*} \sigma_1 \leq \rho \Rightarrow (S_1, \mathcal{A}_2) \tag{2}$$

$$\overline{a} \notin ftv(S_1\sigma_1, S_1\sigma_2) \tag{3}$$

Assume that $fov(S_0\sigma_1, S_0\sigma_2) \subseteq \mathcal{X}$. We know that $fov(S_0(\sigma_2)) = fov(\forall \overline{a}.\rho)$, therefore $fov(S_0\sigma, S_0\rho) \subseteq \mathcal{X}\overline{a}$. By induction hypothesis $fov(S_1\sigma, S_1\rho) \subseteq \mathcal{X}\overline{a}$. But now, because $\overline{a} \notin ftv(S_1\sigma_1)$ and $\overline{a} \notin ftv(S_1(\forall \overline{a}.\rho))$ we have that $fov(S_1\sigma, S_1(\forall \overline{a}.\rho)) \subseteq \mathcal{X}$.

For the second part we have the following cases for the last rule used in the derivation.

- Case ASPEC. Here we have that $\mathcal{A}_0\overline{\beta} \notin vars(S_0) \cup ftv((\forall \overline{a}.\rho_1), \rho_2)$ and $(S_0, \mathcal{A}_0\overline{\beta}) \vdash^{dsk*} \forall \overline{a}.\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$, given that $(S_0, \mathcal{A}_0) \vdash^{dsk*} [\overline{a \mapsto \beta}]\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$. Assume that $fov(S_0(\forall \overline{a}.\rho_1), S_0\rho_2) \subseteq \mathcal{X}$. Then trivially $fov(S_0([\overline{a \mapsto \beta}]\rho_1, S_0\rho_2) \subseteq \mathcal{X}$, since $\overline{\beta} \notin dom(S_0)$. Then we can apply the induction hypothesis to get that $fov(S_1([\overline{a \mapsto \beta}]\rho_1), S_1\rho_2) \subseteq \mathcal{X}$ and therefore $fov(S_1(\forall \overline{a}.\rho_1), S_1\rho_2) \subseteq \mathcal{X}$.

- Case AFUN1 (similarly for AFUN2). We have that $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\rho, \sigma_3, \sigma_4)$ and $(S_0, \mathcal{A}_0) \vdash^{dsk*} \rho \leq \sigma_3 \to \sigma_4 \Rightarrow (S_3, \mathcal{A}_3)$, given that $(S_0, \mathcal{A}_0) \vdash^{\to} \rho = \sigma_1 \to \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$, $(S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_3 \leq \sigma_1 \Rightarrow (S_2, \mathcal{A}_2)$, $(S_1, \mathcal{A}_1) \vdash^{dsk*} \sigma_2 \leq \sigma_4 \Rightarrow (S_3, \mathcal{A}_3)$. Assume that $fov(S_0\rho, S_0\sigma_3, S_0\sigma_4) \subseteq \mathcal{X}$.

  Claim 1: $fov(S_1\rho, S_1\sigma_1, S_1\sigma_2) \subseteq \mathcal{X}$. This follows directly by Lemma 4.10.

  Claim 2: $fov(S_1\sigma_3) \subseteq \mathcal{X}$. Clearly all the ordinary variables of σ_3 are in $\mathcal{X}$, therefore we need to consider the meta variables. Assume $\gamma \in ftv(\sigma_3)$ and assume also that $\gamma \notin ftv(\sigma_1, \sigma_2, \rho)$, because in the last case we are done by the first claim. Now we need to consider what happens in the ordinary variables of $S_1\gamma$. If $\gamma \notin dom(S_1)$ we are trivially okay. If not, then observe that $dom(S_1) \subseteq vars(S_0) \cup ftv(\rho) \cup (\mathcal{A}_0 - \mathcal{A}_1)$ by Lemma 4.12. Then we have two problematic cases.

    - $\gamma \in dom(S_0)$. Then there exists a $\tau = S_0\gamma$. If that type contains only ordinary variables, then we are okay since $S_1\gamma = S_0\gamma$ in that case and by assumptions these ordinary variables are in $\mathcal{X}$. Suppose however that there exists a $\zeta \in \tau$ such that $\zeta \in dom(S_1)$—if it is not in the domain of S_1 we are again trivially okay, since S_1 extends S_0. Then it must be that $\zeta \in ftv(S_0\sigma_1, S_0\sigma_2) \cup (\mathcal{A}_0 - \mathcal{A}_1)$. But it cannot be in $\mathcal{A}_0 - \mathcal{A}_1$ because all variables of S_0 are disjoint from $\mathcal{A}_0$. Then if it is in $ftv(S_0\sigma_1, S_0\sigma_2)$ we also have that $fov(S_1\zeta) \subseteq fov(S_1S_0\sigma_1, S_1S_0\sigma_2)$ and then $fov(S_1\zeta) \subseteq fov(S_1\sigma_1, S_1\sigma_2) \subseteq \mathcal{X}$ by previous claim. therefore $S_1\gamma \subseteq \mathcal{X}$.

– $\gamma \notin dom(S_0)$ which means that $\gamma \in vars(R_1)$ where $R_1 \cdot S_0 = S_1$. But then it must be that $\gamma \in ftv(S_0\rho) \cup (\mathcal{A}_0 - \mathcal{A}_1)$. It cannot be the case that $\gamma \in \mathcal{A}_0 - \mathcal{A}_1$, so $\gamma \in ftv(S_0\rho)$ which means that $fov(S_1\gamma) \in fov(S_1S_0\rho)$ and by idempotency $fov(S_1\gamma) \in fov(S_1\rho \subseteq \mathcal{X}$ because of the first claim.

Then, by induction hypothesis we get that $fov(S_2\sigma_1) \subseteq fov(S_2\sigma_3) \subseteq \mathcal{X}$.

Claim 3: $fov(S_2\sigma_2, S_2\sigma_4) \subseteq \mathcal{X}$. Again, we don't care about the free ordinary variables of σ_2 and σ_4, they are going to be in $\mathcal{X}$ by assumptions. Consider then a $\gamma \in ftv(\sigma_2, \sigma_4)$, such that $\gamma \in dom(S_2)$—if it is not in the domain we are again trivially okay. Then we want to examine what happens to $S_2\gamma$. It must be that $\gamma \in vars(S_1) \cup ftv(\sigma_1, \sigma_3) \cup (\mathcal{A}_1 - \mathcal{A}_2)$. If $\gamma \in ftv(\sigma_1, \sigma_3)$ we are okay. If not, since γ cannot be in $\mathcal{A}_1 - \mathcal{A}_2$, we are left with two cases:

– Here $\gamma \in dom(S_1)$. That means that there exists a $\tau = S_1\gamma$. If τ contains only ordinary variables or meta variables not in the domain of S_2 we are okay. Consider now the case where there exists a $\zeta \in \tau$ such that $\zeta \in dom(S_2)$. Then, ζ is in the extension of S_1), therefore $\zeta \in ftv(S_1\sigma_1, S_1\sigma_3) \cup (\mathcal{A}_1 - \mathcal{A}_2)$. But it can't be in $\mathcal{A}_1 - \mathcal{A}_2$ as it is also in S_1 and the variables of S_1 are disjoint from $\mathcal{A}_1$. Then it must be in $ftv(S_1\sigma_1, S_1\sigma_3)$ therefore $fov(S_2\zeta) \subseteq \mathcal{X}$, since $fov(S_2\zeta) \subseteq fov(S_2S_1\sigma_1, S_2S_1\sigma_3)$ and $fov(S_2S_1\sigma_1, S_2S_1\sigma_3) = fov(S_2\sigma_1, S_2\sigma_3) \subseteq \mathcal{X}$.

– Here $\gamma \notin dom(S_1)$. This means that $\gamma \in vars(R_2)$ where $R_2 \cdot S_1 = S_2$. Then we know that $vars(R_2) \subseteq ftv(S_1\sigma_3, S_1\sigma_1) \cup (\mathcal{A}_1 - \mathcal{A}_2)$, but $\gamma \notin \mathcal{A}_1$, therefore $\gamma \in ftv(S_1\sigma_3, S_1\sigma_1)$. Then $fov(S_2\gamma) \in fov(S_2S_1\sigma_3, S_2S_1\sigma_1)$, or $fov(S_2\gamma) \in fov(S_2\sigma_3, S_2\sigma_1) \subseteq \mathcal{X}$.

Now we can apply the induction hypothesis again to get that $fov(S_3\sigma_2, S_3\sigma_4) \subseteq \mathcal{X}$. At this point let us consider again what we have established so far and what we want to prove. We want to prove the following:

$$
\begin{aligned}
fov(S_3\rho) &\subseteq \mathcal{X} \\
fov(S_3\sigma_3) &\subseteq \mathcal{X} \\
fov(S_3\sigma_4) &\subseteq \mathcal{X}
\end{aligned}
$$

But it is easily derivable from Lemma 4.9 and Lemma 4.12 that $S_3\rho = S_3\sigma_1 \to S_3\sigma_2$. But we already have that $fov(S_3\sigma_2, S_3\sigma_4) \subseteq \mathcal{X}$. Therefore we only need to establish that $fov(S_3\sigma_1, S_3\sigma_3) \subseteq \mathcal{X}$.

Claim 4: $fov(S_3\sigma_1, S_3\sigma_3) \subseteq \mathcal{X}$. Again, we don't care about the free ordinary variables of σ_1 and σ_3; these are in $\mathcal{X}$ by assumptions. Consider $\gamma \in ftv(\sigma_3, \sigma_1)$ and in fact let $\gamma \notin ftv(\sigma_2, \sigma_4)$—otherwise we are ok. Also suppose that $\gamma \in dom(S_3)$ otherwise we are trivially okay. Then, it must be that $\gamma \in vars(S_2) \cup ftv(\sigma_2, \sigma_4) \cup (\mathcal{A}_2 - \mathcal{A}_3)$. There are two non-trivial cases.

– $\gamma \in dom(S_2)$. Then let $\tau = S_2\gamma$. If τ does not contain meta variables, or contains meta variables not in the domain of S_3 we are okay. Consider the bad case where there exists a $\zeta \in \tau$ such that $\zeta \in dom(S_3)$. Then $\zeta \in ftv(S_2\sigma_2, S_2\sigma_4) \cup (\mathcal{A}_2 - \mathcal{A}_3)$. It can't be in $\mathcal{A}_2 - \mathcal{A}_3$ because it is also in S_2 and S_2 does not contain variables from $\mathcal{A}_2$. So it must be in $ftv(S_2\sigma_2, S_2\sigma_4)$, therefore $fov(S_3\zeta) \subseteq \mathcal{X}$, since $fov(S_3\zeta) \subseteq ftv(S_3S_2\sigma_2, S_3S_2\sigma_4)$. Then also $S_3\gamma \subseteq \mathcal{X}$.

– $\gamma \notin dom(S_2)$. Then it must be the case that $\gamma \in R_3$, where $R_3 \cdot S_2 = S_3$. Then also $vars(R_3) \subseteq ftv(S_2\sigma_2, S_2\sigma_4) \cup (\mathcal{A}_2 - \mathcal{A}_3)$, but we know that $\gamma \notin \mathcal{A}_2$. Then $fov(S_3\gamma) \subseteq fov(S_3S_2\sigma_2, S_3S_2\sigma_4)$, or $fov(S_3\gamma) \subseteq fov(S_3\sigma_2, S_3\sigma_4) \subseteq \mathcal{X}$.

• Case AMONO. Follows directly by Lemma 4.7.

□

Now two lemmas about the variables during inference and checking.

Lemma 4.14.

1. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ $\Rightarrow$ $\exists R$ *such that* $S_1 = R \cdot S_0$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$ $\qquad \mathit{vars}(S_1) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{ftv}(\rho) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{vars}(R) \subseteq \mathit{ftv}(S_0\Gamma) \cup (\mathcal{A}_0 - \mathcal{A}_1)$

2. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup \mathit{ftv}(\rho)$ $\Rightarrow$ $\exists R$ *such that* $S_1 = R \cdot S_0$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Downarrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$ $\qquad \mathit{vars}(S_1) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{ftv}(\rho) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{vars}(R) \subseteq \mathit{ftv}(S_0\Gamma) \cup \mathit{ftv}(S_0\rho) \cup (\mathcal{A}_0 - \mathcal{A}_1)$

3. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ $\Rightarrow$ $\exists R$ *such that* $S_1 = R \cdot S_0$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$ $\qquad \mathit{vars}(S_1) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{ftv}(\sigma) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{vars}(R) \subseteq \mathit{ftv}(S_0\Gamma) \cup (\mathcal{A}_0 - \mathcal{A}_1)$

4. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup \mathit{ftv}(\sigma)$ $\Rightarrow$ $\exists R$ *such that* $S_1 = R \cdot S_0$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Downarrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$ $\qquad \mathit{vars}(S_1) \subseteq \mathit{ftv}(\Gamma) \cup \mathit{ftv}(\sigma) \cup \mathit{vars}(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$
    $\qquad \mathit{vars}(R) \subseteq \mathit{ftv}(S_0\Gamma) \cup \mathit{ftv}(S_0\sigma) \cup (\mathcal{A}_0 - \mathcal{A}_1)$

Proof. Straightforward induction on the derivations. $\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Lemma 4.15.

1. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ $\Rightarrow$ $\mathit{fov}(S_1\Gamma) \cup \mathit{fov}(S_1\rho) \subseteq \mathcal{X}$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
    $\mathit{fov}(S_0\Gamma) \subseteq \mathcal{X}$

2. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup \mathit{ftv}(\rho)$ $\Rightarrow$ $\mathit{fov}(S_1\Gamma) \cup \mathit{fov}(S_1\rho) \subseteq \mathcal{X}$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Downarrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
    $\mathit{fov}(S_0\Gamma) \cup \mathit{fov}(S_0\rho) \subseteq \mathcal{X}$

3. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ $\Rightarrow$ $\mathit{fov}(S_1\Gamma) \cup \mathit{fov}(S_1\sigma) \subseteq \mathcal{X}$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$
    $\mathit{fov}(S_0\Gamma) \subseteq \mathcal{X}$

4. $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0) \cup \mathit{ftv}(\sigma)$ $\Rightarrow$ $\mathit{fov}(S_1\Gamma) \cup \mathit{fov}(S_1\sigma) \subseteq \mathcal{X}$
    $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Downarrow^{poly} t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
    $\mathit{fov}(S_0\Gamma) \cup \mathit{fov}(S_0\sigma) \subseteq \mathcal{X}$

Proof. We prove the four goals simultaneously by induction on the algorithmic relations. We are going to use the results of Lemma 4.14 as well as Lemma 4.13 extensively.

For the first part we consider the following cases.

- Case ALG-VAR. We have that $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ and $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow x : \rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $x : \sigma \in \Gamma$ and $(S_0, \mathcal{A}_0) \vdash_\Uparrow^{inst} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)$. Assume that $\mathit{fov}(S_0\Gamma) \subseteq \mathcal{X}$. Then by an inversion we see that $S_1 = S_0$ and $\mathit{fov}(S_1\rho) \subseteq \mathit{fov}(S_1\sigma)$, because $\mathcal{A}_0 - \mathcal{A}_1$ is going to be a set of meta variables not in the variables of S_0.

- Case ALG-ABS1. Here let $\mathcal{A}_0\beta \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ and $(S_0, \mathcal{A}_0\beta)\,; \Gamma \vdash_\Uparrow \backslash x.\,t : \beta \to \rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $(S_0, \mathcal{A}_0)\,; \Gamma, x : \beta \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$. Assume $\mathit{fov}(S_0\Gamma) \subseteq \mathcal{X}$. Then also $\mathit{fov}(S_0(\Gamma, x : \beta)) \subseteq \mathcal{X}$, since $\beta \notin \mathit{vars}(S_0)$. Then, by induction hypothesis $\mathit{fov}(S_1\Gamma, S_1\beta) \subseteq \mathcal{X}$ and $\mathit{fov}(S_1\rho) \subseteq \mathcal{X}$. But that is exactly what is required for this case.

- Case ALG-AABS1. Assume that $\mathcal{A}_0 \notin \mathit{ftv}(\Gamma) \cup \mathit{vars}(S_0)$ and $(S_0, \mathcal{A}_0)\,; \Gamma \vdash_\Uparrow \backslash x::\sigma.\,t : \sigma \to \rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $(S_0, \mathcal{A}_0)\,; \Gamma, x : \sigma \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$. Assume $\mathit{fov}(S_0\Gamma) \subseteq \mathcal{X}$. Then also $\mathit{fov}(S_0(\Gamma, x : \sigma)) \subseteq \mathcal{X}$, since annotations are closed. Then, by induction hypothesis $\mathit{fov}(S_1\Gamma, S_1\sigma) \subseteq \mathcal{X}$ and $\mathit{fov}(S_1\rho) \subseteq \mathcal{X}$ and we are done.

- Case ALG-APP. Here we have that $\mathcal{A}_0 \notin vars(S_0) \cup ftv(\Gamma)$ and $(S_0, \mathcal{A}_0)\, ;\Gamma \vdash_{\Uparrow} t\ u : \rho \Rightarrow (S_4, \mathcal{A}_4)$, given that

$$(S_0, \mathcal{A}_0)\, ;\Gamma \vdash_{\Uparrow} t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1) \tag{1}$$

$$(S_1, \mathcal{A}_1) \vdash^{\rightarrow} \rho_1 = \sigma \rightarrow \sigma' \Rightarrow (S_2, \mathcal{A}_2) \tag{2}$$

$$(S_2, \mathcal{A}_2)\, ;\Gamma \vdash_{\Downarrow}^{poly} u : \sigma \Rightarrow (S_3, \mathcal{A}_3) \tag{3}$$

$$(S_3, \mathcal{A}_3) \vdash_{\Uparrow}^{inst} \sigma' \le \rho \Rightarrow (S_4, \mathcal{A}_4) \tag{4}$$

Let us assume that $fov(S_0\Gamma) \subseteq \mathcal{X}$. Then by induction we get that $fov(S_1\rho_1) \subseteq \mathcal{X}$ and $fov(S_1\Gamma) \subseteq \mathcal{X}$. Now, by the arrow unification variables lemma we get that $fov(S_2\sigma, S_2\sigma', S_2\rho_1) \subseteq \mathcal{X}$. At this point we need to show that $fov(S_2\Gamma) \subseteq \mathcal{X}$ to be able to apply the induction hypothesis further.

Claim 1: $fov(S_2\Gamma) \subseteq \mathcal{X}$. First, all the free ordinary variables of Γ are by assumptions in $\mathcal{X}$. Then we are interested to see what happens to the meta variables of Γ and in particular those that are in the domain of S_2. Consider $\gamma \in dom(S_2)$ such that $\gamma \in ftv(\Gamma)$. Then by previous lemmas we know that $\gamma \in vars(S_1) \cup ftv(\rho_1, \sigma, \sigma') \cup (\mathcal{A}_1 - \mathcal{A}_2)$. If $\gamma \in ftv(\rho_1, \sigma, \sigma')$ then we are okay. Also $\gamma \notin \mathcal{A}_1 - \mathcal{A}_2$ since it is a free variable of Γ. Then we have two cases.

- Suppose that $\gamma \in dom(S_1)$, that is, there exists a type $\tau = S_1\gamma$. If τ does not contain meta variables we are okay, since $S_2\gamma = \S_1\gamma = \tau$ in this case. However suppse that there exists a variable $\zeta \in \tau$. If that ζ is not in the domain of S_2 then we are okay. If however $\zeta \in dom(S_2)$ then it is in the extension of S_1, therefore we get that $\zeta \in ftv(S_1\rho_1) \cup (\mathcal{A}_1 - \mathcal{A}_2)$. But it cannot be in in $\mathcal{A}_1 - \mathcal{A}_2$ since we know that $vars(S_1) \notin \mathcal{A}_1$ by previous lemma. Then $fov(S_2\zeta) \in fov(S_2S_1\rho_1)$ or $fov(S_2\zeta) \in fov(S_2\rho_1) \subseteq \mathcal{X}$. Then also $fov(S_2\gamma) \subseteq \mathcal{X}$ since S_2 is an extension of S_1.
- Suppose that $\gamma \notin dom(S_1)$, then $\gamma \in vars(R_2)$ where $S_2 = R_2 \cdot S_1$, therefore $\gamma \in ftv(S_1\rho_1) \cup (\mathcal{A}_1 - \mathcal{A}_2)$. But it cannot be in $\mathcal{A}_1 - \mathcal{A}_2$ therefore $fov(S_2\gamma) \subseteq fov(S_2S_1\rho_1) = fov(S_2\rho_1) \subseteq \mathcal{X}$.

Now that we proved this claim, it is easy to confirm that the conditions are appropriate to apply the induction hypothesis to get that $fov(S_3\Gamma) \subseteq \mathcal{X}$ and $fov(S_3\sigma) \subseteq \mathcal{X}$. Now with a simple inversion we get that $S_3 = S_4$ and $fov(S_3\rho) = fov(S_3\sigma')$. Then to finish the case we need to show that $fov(S_3\sigma') \subseteq \mathcal{X}$.

Claim 2: $fov(S_3\sigma') \subseteq \mathcal{X}$. The claim uses a similar proof technique as the previous claim.

- Case ALG-LET. Suppose $\mathcal{A}_0 \notin ftv(\Gamma) \cup ftv(S_0)$ and we have that $(S_0, \mathcal{A}_0)\, ;\Gamma \vdash_{\Uparrow} \texttt{let } x = u \texttt{ in } t : \rho \Rightarrow (S_2, \mathcal{A}_2)$, given that $S_0, \mathcal{A}_0)\, ;\Gamma \vdash_{\Uparrow}^{poly} u : \sigma \Rightarrow (S_1, \mathcal{A}_1)$ and $S_1, \mathcal{A}_1)\, ;\Gamma, x : \sigma \vdash_{\Uparrow}^{poly} t : \rho \Rightarrow (S_2, \mathcal{A}_2)$. Assume that $fov(S_0\Gamma) \subseteq \mathcal{X}$. Then by induction we get that $fov(S_1\Gamma) \subseteq \mathcal{X}$ and $fov(S_1\sigma) \subseteq \mathcal{X}$. Then it is easy to confirm using previous lemmas that $\mathcal{A}_1 \notin ftv(\Gamma) \cup vars(S_1)$ and we can apply again the induction hypothesis to get that $fov(S_2\Gamma, S_2\sigma) \subseteq \mathcal{X}$ and $fov(S_2\rho) \subseteq \mathcal{X}$ as required for this case.

The second part follows the structure of the first part using the technique shown in the claims above to establish the appropriate conditions at each step.

For the third part, we have that $\mathcal{A}_0\overline{b} \notin ftv(\Gamma) \cup vars(S_0)$ and $(S_0, \mathcal{A}_0\overline{b}); \Gamma \vdash_{\Uparrow}^{poly} t : \forall \overline{b}.\,[\overline{\alpha \mapsto b}]S_1\rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $(S_0, \mathcal{A}_0)\, ;\Gamma \vdash_{\Uparrow} t : \rho \Rightarrow (S_1, \mathcal{A}_1)$ and $\overline{\alpha}$ are the meta variables of $ftv(S_1\rho) - ftv(S_1\Gamma)$. Assume that $fov(S_0\Gamma) \subseteq \mathcal{X}$. Then we can apply the induction hypothesis to get that $fov(S_1\Gamma) \subseteq \mathcal{X}$ and $fov(S_1\rho) \subseteq \mathcal{X}$. But then it must also be $fov(\forall \overline{b}.\,[\overline{\alpha \mapsto b}]S_1\rho) \subseteq \mathcal{X}$ and this case is done.

For the fourth part, we have that $\mathcal{A}_0\overline{b} \notin ftv(\Gamma) \cup vars(S_0) \cup ftv(\sigma)$ and $(S_0, \mathcal{A}_0); \Gamma \vdash_{\Downarrow}^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_2)$, given that $\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall \overline{a}.\rho \Rightarrow \mathcal{A}_1$, $(S_0, \mathcal{A}_1); \Gamma \vdash_{\Downarrow} t : \rho \Rightarrow (S_1, \mathcal{A}_2)$, and $\overline{a} \notin ftv(S_1\Gamma, S_1\sigma)$. Assume that $fov(S_0\Gamma) \cup fov(S_0(\sigma)) \subseteq \mathcal{X}$ or equivalently $fov(S_0\Gamma) \cup fov(S_0(\forall \overline{a}.\rho)) \subseteq \mathcal{X}$. Therefore $fov(S_0\Gamma) \cup fov(S_0\rho) \subseteq \mathcal{X}\overline{a}$. Then, by induction hypothesis $fov(S_1\Gamma) \cup fov(S_1\rho) \subseteq \mathcal{X}\overline{a}$. But now, since $\overline{a} \notin ftv(S_1\Gamma)$ and $\overline{a} \notin ftv(S_1(\sigma))$, it must be that $fov(S_1\Gamma) \cup fov(S_1\sigma) \subseteq \mathcal{X}$. $\qquad\square$

Lemma 4.16 (Weak Prenex Conversion Completeness). *If $pr(S\sigma) = \forall\overline{a}.\rho_a$ then $\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall\overline{b}.\rho_b \Rightarrow \mathcal{A}_1$, such that $S(\forall\overline{b}.\rho_b) = \forall\overline{a}.\rho_a$.*

Proof. Easy induction on σ. Moreover by Lemma 4.11 it is also the case that $\overline{b} \subseteq \mathcal{A}_0 - \mathcal{A}_1$. $\qquad\square$

Now a completeness result for the algorithmic subsumption relation.

Lemma 4.17 (Algorithmic Subsumption Completeness). *Suppose we are given unifiers S, S_0, a context Γ and two polytypes σ_1 and σ_2. Then*

1. *If $\vdash^{dsk} SS_0\sigma_1 \leq SS_0\sigma_2$ is canonical then $\forall\mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \sigma_2)$ we have that $(\mathcal{A}, S_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, and $vars(R) \notin \mathcal{A}_1$, $vars(S) \notin \mathcal{A}_1$.*

2. *If $\vdash^{dsk} SS_0\sigma_1 \leq SS_0\sigma_2$ is prenex-canonical then $\forall\mathcal{A}_0 \notin vars(S, S_0)\cup ftv(\sigma_1, \sigma_2)$ we have that $(\mathcal{A}, S_0) \vdash^{dsk*} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, and $vars(R) \notin \mathcal{A}_1$, $vars(S) \notin \mathcal{A}_1$.*

As a corollary, if $\vdash^{dsk} SS_0\sigma_1 \leq SS_0\sigma_2$ $\forall\mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \sigma_2)$ we have that $(\mathcal{A}, S_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, and $vars(R) \notin \mathcal{A}_1$, $vars(S) \notin \mathcal{A}_1$.

Proof. We prove the two goals simultaneously by induction on the height of the derivations. We proceed with the last rule used.

Part 1: For this part the only rule that could have been used is rule SKOL. For this case we have that $\vdash^{dsk} SS_0\sigma_1 \leq SS_0\sigma_2$, given that

$$pr(SS_0\sigma_2) = \forall\overline{a}.\rho_a \tag{1}$$

$$\overline{a} \notin ftv(SS_0\sigma_1) \tag{2}$$

$$\vdash^{dsk} SS_0\sigma_1 \leq \rho_a \tag{3}$$

Consider an appropriate symbol supply $\mathcal{A}_0$, such that $\mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \sigma_2)$. By Lemma 4.16 we have that $\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall\overline{b}.\rho_b \Rightarrow \mathcal{A}_1$ such that $SS_0(\forall\overline{b}.\rho_b) = \forall\overline{a}.\rho_a$. By Lemma 4.11 $\overline{b} \in \mathcal{A}_0$. This means that $SS_0\rho_b = \overline{[a \mapsto b]}\rho_a$. From this, the substitution lemma and (2), and (3) we get that $\vdash^{dsk} SS_0\sigma_1 \leq SS_0\rho_b$. Moreover this last derivation is prenex-canonical, as (3). Therefore by induction $(S_0, \mathcal{A}_1) \vdash^{dsk*} \sigma_1 \leq \rho_b \Rightarrow (S_1, \mathcal{A}_2)$. and $\exists R$ such that $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_1 - \mathcal{A}_2}$, since $S_0 = S_1$. Equivalently $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_2}$. Moreover $vars(R) \notin \mathcal{A}_2$ and $vars(S_1) \notin \mathcal{A}_2$. Finally to be able to apply the rule ASKOL we need to show that $\overline{b} \notin ftv(S_1\sigma_1, S_1\sigma_2)$. Assume by contradiction that for some $b \in \overline{b}$ it is the case that $b \in ftv(S_1\sigma_1, S_1\sigma_2)$. This means that there exists a $\gamma \in ftv(\sigma_1, \sigma_2)$ and $b \in S_1\gamma$ because $b \notin ftv(\sigma_1, \sigma_2)$. Moreover by the freshness conditions $\gamma \notin \mathcal{A}_0$ and therefore Then $SS_0\gamma = RS_1\gamma$. But then it must be $b \in vars(S, S_0)$, a contradiction.

Part 2: For this part we notice that a prenex-canonical derivation ends with a trivial SKOL application which can be ommited; therefore we have to examine all other rules.

- Case MONO. In this case since $SS_0\sigma_1$ and $SS_0\sigma_2$ are monotypes, it must be that $\sigma_1 = \tau_1$ and $\sigma_2 = \tau_2$ for some monotypes τ_1 and τ_2. Pick an arbitrary $\mathcal{A}_0$ that satisfies the freshness conditions, that is $\mathcal{A}_0 \notin ftv(\sigma_1)$, $\mathcal{A}_0 \notin ftv(\sigma_2)$, $\mathcal{A}_0 \notin vars(S) \cup vars(S_0)$. By Lemma 4.8 we have that $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$ and $S \cdot S_0 = R \cdot S_1$ for some R. Moreover $vars(R) \subseteq vars(S) \cup vars(S_0) \cup vars(S_1)$ which means $vars(R) \notin \mathcal{A}_0$. Finally $vars(S_1) \subseteq vars(S_0) \cup ftv(\tau_1, \tau_2)$, again disjoint from $\mathcal{A}_0$. Then by applying the rule AMONO we are done.

- Case FUN. In this case we have that $SS_0\sigma_1 = \sigma'_{11} \to \sigma'_{12}$ and $SS_0\sigma_2 = \sigma'_{21} \to \sigma'_{22}$. Now, it must be the case that $\sigma_1 = \rho_1$ and $\sigma_2 = \rho_2$ for some ρ_1 and ρ_2 because the substitution $S \cdot S_0$ returns back arrow types and not quantified types. Also, it cannot be the case that both ρ_1 and ρ_2 are monotypes, because then we would be in the MONO case. We split in cases depending on which of the two types is a ρ arrow type.

  Assume that $\rho_1 = \sigma_{11} \to \sigma_{12}$, and pick a supply $\mathcal{A}_0$ such that $\mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \sigma_2)$. Then it must be that $\vdash^{dsk} SS_0\sigma_{11} \to SS_0\sigma_{12} \le SS_0\rho_2$. Since $SS_0\rho_2 = \sigma'_{21} \to \sigma'_{22}$ by unification completeness Lemma 4.9 (it is easy to confirm that the freshness conditions for $\mathcal{A}_0$ are sufficient) we have that $(S_0, \mathcal{A}_0) \vdash^{\to} \rho_2 = \sigma_{21} \to \sigma_{22} \Rightarrow (S_1, \mathcal{A}_1)$ such that $\exists R_u$ with $S \cdot S_0 = R_u \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$ and $R_u S_1 \sigma_{21} = \sigma'_{21}$ and $R_u S_1 \sigma_{22} = \sigma'_{22}$. Moreover $vars(R_u) \notin \mathcal{A}_1$ and $vars(S_1) \notin \mathcal{A}_1$. Then by the premises of the rule $\vdash^{dsk} \sigma'_{21} \le \sigma_{11}$ or $\vdash^{dsk} R_u S_1 \sigma_{21} \le SS_0\sigma_{11}$ or $\vdash^{dsk} R_u S_1 \sigma_{21} \le R_u S_1 \sigma_{11}$ since $ftv(\sigma_{11}) \notin (\mathcal{A}_0 - \mathcal{A}_1)$. Then we can apply the induction hypothesis for the supply $\mathcal{A}_1$ to get that $(S_1, \mathcal{A}_1) \vdash^{dsk*} \sigma_{21} \le \sigma_{11} \Rightarrow (S_2, \mathcal{A}_2)$ and $\exists R_1$ with $R_u \cdot S_1 = R_1 \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$ which implies that $S \cdot S_0 = R_1 \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$. Moreover $vars(R_1) \notin \mathcal{A}_2$ and $vars(S_2) \notin \mathcal{A}_2$. Now we also know by the premises of the rule that $\vdash^{dsk} \sigma'_{12} \le \sigma'_{22}$ or $\vdash^{dsk} R_1 S_2 \sigma_{12} \le R_1 S_2 \sigma_{22}$. Then it is easy to confirm that the freshness conditions hold for $\mathcal{A}_2$ to apply the induction hypothesis and get that $(\mathcal{A}_2, S_2) \vdash^{dsk} \sigma_{12} \le \sigma_{22} \Rightarrow (\mathcal{A}_3, S_3)$ and $\exists R$ with $R_1 \cdot S_2 = R \cdot S_3 \backslash_{\mathcal{A}_2 - \mathcal{A}_3}$ which gives us that $S \cdot S_0 = R_2 \cdot S_3 \backslash_{\mathcal{A}_0 - \mathcal{A}_3}$. Moreover $vars(R) \notin \mathcal{A}_3$ and $vars(S_3) \notin \mathcal{A}_3$. Then by applying rule AFUN1 and picking the same R we are done.

  The case where $\rho_2 = \sigma_{21} \to \sigma_{22}$ is similar.

- Case SPEC. Here it must be that σ_1 is a polytype and assume that $\sigma_1 = \forall \overline{a}.\rho_1$ and $\sigma_2 = \rho_2$. Assume also without loss of generality that $\overline{a} \notin vars(S, S_0)$. Then we have that $\vdash^{dsk} \forall \overline{a}.SS_0\rho_1 \le \rho_2$ given that $\vdash^{dsk} [\overline{a \mapsto \tau}]SS_0\rho_1 \le SS_0\rho_2$, $\overline{a} \notin \overline{\tau}$. Then consider an arbitrary supply $\mathcal{A}_0\overline{\beta}$ such that $\mathcal{A}_0\overline{\beta} \notin vars(S, S_0) \cup ftv(\sigma_1, \rho_2)$ and rewrite the last derivation as: $\vdash^{dsk} [\overline{\beta \mapsto \tau}]SS_0[\overline{a \mapsto \beta}]\rho_1 \le SS_0\rho_2$, Now because of the freshness conditions this is equivalent to $\vdash^{dsk} [\overline{\beta \mapsto \tau}]SS_0\rho_1 \le [\overline{\beta \mapsto \tau}]SS_0\rho_2$.

  Now we need to be able to apply the induction hypothesis for $\mathcal{A}_0$ but we cannot do this directly because $\overline{\tau}$ might contain variables in $\mathcal{A}_0$. Instead we do the following: Separate the free variables of $\overline{\tau}$ in two sets. Let $\mathcal{X}_1 = ftv(\overline{\tau}) \cap (ftv(SS_0\rho_2, SS_0\rho_1))$ and $\mathcal{X}_2 = ftv(\overline{\tau}) - \mathcal{X}_1$. We know by our assumptions that $\mathcal{A}_0 \notin \mathcal{X}_1$, so the problematic set is $\mathcal{X}_2$. But simply consider a renaming substitution Q from $\mathcal{X}_2$ to a set $\mathcal{X}$ of variables disjoint from $\mathcal{A}_0$. By the substitution lemma, Lemma 2.38, we get that $\vdash^{dsk} [\overline{\beta \mapsto Q\tau}]SS_0\rho_1 \le [\overline{\beta \mapsto Q\tau}]SS_0\rho_2$. And now we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) \vdash^{dsk} [\overline{a \mapsto \beta}]\rho_1 \le \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $[\overline{\beta \mapsto Q\tau}] \cdot S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$ which means that $S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0\overline{\beta} - \mathcal{A}_1}$. We also get that $vars(S_1) \notin \mathcal{A}_1$, $vars(R) \notin \mathcal{A}_1$. Then the rule ASPEC is applicable and taking the same R finishes the case.

$\square$

Next an auxiliary corollary for algorithmic instantiation.
Corollary 4.18 (Algorithmic Instantiation Completeness).

1. $\vdash_{\Downarrow}^{inst} SS_0\sigma_1 \le SS_0\rho_2 \quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \rho_2)$
   $(\mathcal{A}, S_0) \vdash_{\Downarrow}^{inst} \sigma_1 \le \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \, s.t. \ S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$
   $and \ vars(R) \notin \mathcal{A}_1, vars(S) \notin \mathcal{A}_1$

2. $\vdash_{\Uparrow}^{inst} SS_0\sigma_1 \le \rho_2 \quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma_1, \rho_2)$
   $(\mathcal{A}, S_0) \vdash_{\Uparrow}^{inst} \sigma_1 \le \rho'_2 \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \, s.t. \ S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$
   $and \ RS_1\rho'_2 = \rho_2$
   $and \ vars(R) \notin \mathcal{A}_1, vars(S) \notin \mathcal{A}_1, ftv(\rho'_2) \notin \mathcal{A}_1$

Proof. The first part easily follows from Lemma 4.17. For the second part it must be that σ_1 is a polytype $\sigma_1 = \forall \overline{a}.\rho_1$ and assume without loss of generality that $\overline{a} \notin vars(S, S_0)$. Then, using the rule INST we have that $\vdash_\Uparrow^{inst} \forall \overline{a}.SS_0\rho_1 \leq \rho_2$ given that $\vdash_\Uparrow^{inst\rho} [\overline{a \mapsto \tau}]SS_0\rho_1 \leq \rho_2$ and this with an extra inversion using rule INST1 gives us that $[\overline{a \mapsto \tau}]SS_0\rho_1 = \rho_2$ for that $\overline{\tau}$. Now the algorithm, using rule AINST will give us that for an appropriate $\mathcal{A}_0\overline{\beta}$ we have that $(S_0, \mathcal{A}_0\overline{\beta}) \vdash_\Uparrow^{inst} \forall \overline{a}.\rho_1 \leq \rho_2' \Rightarrow (S_1, \mathcal{A}_1)$ when $(S, \mathcal{A}_0) \vdash_\Uparrow^{inst\rho} [\overline{a \mapsto \beta}]\rho_1 \leq \rho_2' \Rightarrow (S_1, \mathcal{A}_1)$. Now using the rule IINSTRHO we see that $\rho_2' = [\overline{a \mapsto \beta}]\rho_1$, $S_1 = S_0$. Take $R = [\overline{\beta \mapsto \tau}] \cdot S$. Then we have that $S \cdot S_0 = R \cdot S_1\backslash_{\overline{\beta}}$ as required. Moreover the variable freshness conditions are also satisfied. Finally we need to show that $RS_1\rho_2' = \rho_2$. We have that

$$
\begin{aligned}
RS_1\rho_2' &= [\overline{\beta \mapsto \tau}]SS_0[\overline{a \mapsto \beta}]\rho_1 \\
&= [\overline{\beta \mapsto \tau}][\overline{a \mapsto \beta}]SS_0\rho_1 \\
&= [\overline{a \mapsto \tau}]SS_0\rho_1 \\
&= \rho_2
\end{aligned}
$$

$\square$

Now the main completeness result.

Lemma 4.19 (Algorithmic Completeness).

1. $SS_0\Gamma \vdash_\Uparrow t : \rho$ $\quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\Gamma) \cup ftv(\rho)$
   $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow t : \rho' \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \ s.t. \ S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1} \ and \ RS_1\rho' = \rho$
   $ftv(\rho') \notin \mathcal{A}_1, vars(R) \notin \mathcal{A}_1, vars(S_1) \notin \mathcal{A}_1$

2. $SS_0\Gamma \vdash_\Downarrow t : SS_0\rho$ $\quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\rho) \cup ftv(\Gamma).$
   $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Downarrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \ s.t. \ S \cdot S_0 = R \cdot S\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$
   $vars(R) \notin \mathcal{A}_1, vars(S_1) \notin \mathcal{A}_1$

3. $SS_0\Gamma \vdash_\Uparrow^{poly} t : \sigma$ $\quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\Gamma) \cup ftv(\sigma)$
   $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow^{poly} t : \sigma' \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \ s.t. \ S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1} \ and \ \vdash^{dsk} RS_1\sigma' \leq_{sh} \sigma$
   $ftv(\sigma') \notin \mathcal{A}_1, vars(R) \notin \mathcal{A}_1, vars(S_1) \notin \mathcal{A}_1$

4. $SS_0\Gamma \vdash_\Downarrow^{poly} t : SS_0\sigma$ $\quad \Rightarrow \quad \forall \mathcal{A}_0 \notin vars(S, S_0) \cup ftv(\sigma) \cup ftv(\Gamma)$
   $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Downarrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$
   $\exists R \ s.t. \ S \cdot S_0 = R \cdot S\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$
   $vars(R) \notin \mathcal{A}_1, vars(S_1) \notin \mathcal{A}_1$

Proof. We prove simultaneously all goals, by induction on the height of the derivations. For each case the induction hypothesis asserts that all goals hold for derivations of smaller heights. We proceed by case analysis on the last rules used.

For the first part we have the following cases.

- Case VAR. The result follows easily from Corollary 4.18.

- Case ABS1. In this case we have that $SS_0\Gamma \vdash_\Uparrow \backslash x.t : \tau \to \rho$, given that $SS_0\Gamma, x : \tau \vdash_\Uparrow t : \rho$. Consider a symbol supply $\mathcal{A}_0\beta$ that satisfies the freshness conditions. This is the same as writing $[\beta \mapsto \tau]SS_0\Gamma, x : \beta \vdash_\Uparrow t : \rho$, because of the freshness conditions. Then we can apply the induction hypothesis with $\mathcal{A}_0$ to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow t : \rho' \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $[\beta \mapsto \tau]S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, $RS_1\rho' = \rho$ and $ftv(\rho') \notin \mathcal{A}_1$, $ftv(S_1) \notin \mathcal{A}_1$, $ftv(R) \notin \mathcal{A}_1$. Then we can apply the rule ALG-ABS1 to get that $(S_0, \mathcal{A}_0\beta); \Gamma \vdash \backslash x.t : \beta \to \rho' \Rightarrow (S_1, \mathcal{A}_1)$. But then $RS_1(\beta \to \rho') = RS_1\beta \to RS_1\rho' = \tau \to \rho$ as required and $S \cdot S_0 = R \cdot S_1\backslash_{\mathcal{A}_0\beta - \mathcal{A}_1}$.

- Case AABS1. Here we have that $SS_0\Gamma \vdash_\Uparrow \backslash x :: \sigma . t : \sigma \to \rho$ given that $SS_0\Gamma, x : \sigma \vdash_\Uparrow t : \rho$. And consider any supply $\mathcal{A}_0$ with the appopriate freshness conditions. Because the type annotations are closed we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) ; \Gamma, x : \sigma \vdash_\Uparrow t : \rho' \Rightarrow (S_1, \mathcal{A}_1)$ and there exists an R such that $S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, $RS_1\rho' = \rho$, $vars(R) \notin \mathcal{A}_1$, $vars(S) \notin \mathcal{A}_1$, $ftv(\rho') \notin \mathcal{A}_1$. Finally we can apply rule ALG-AABS1 to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow \backslash x :: \sigma . t : \sigma \to \rho' \Rightarrow (S_1, \mathcal{A}_1)$ and $RS_1(\sigma \to \rho') = \sigma \to RS_1\rho' = \sigma \to \rho$ as required, since annotations are closed.

- Case APP. In this case we have $SS_0\Gamma \vdash_\Uparrow t\, u : \rho$, given that $SS_0\Gamma \vdash_\Uparrow t : \sigma \to \sigma'$, $SS_0\Gamma \vdash_\Downarrow^{poly} u : \sigma$ and $\vdash_\Uparrow^{inst} \sigma' \le \rho$. Consider any symbol supply $\mathcal{A}_0$ that satisfies the freshness conditions, that is, $\mathcal{A}_0 \notin vars(S, S_0)$, $\mathcal{A}_0 \notin ftv(\rho)$, $\mathcal{A}_0 \notin ftv(\Gamma)$. Then let us consider the set of variables $\mathcal{X} = ftv(\sigma \to \sigma') - (vars(S, S_0) \cup ftv(\rho) \cup ftv(\Gamma))$ and let Q be a renaming substitution taking $\mathcal{X}$ to set of variables disjoint from $\mathcal{A}_0$. Then by the substitution lemma we can create derivations: $SS_0\Gamma \vdash_\Uparrow t : Q\sigma \to Q\sigma'$, $SS_0\Gamma \vdash_\Downarrow^{poly} u : Q\sigma$ and $\vdash^{inst} Q\sigma' \le \rho$, and these derivations have the same height as the original derivations. Now we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1)$ such that $\exists R_1$ with $S \cdot S_0 = R_1 \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, $vars(R_1) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\rho_1) \notin \mathcal{A}_1$. $R_1 S_1 \rho_1 = Q\sigma \to Q\sigma'$. Now it is easy to verify that $\mathcal{A}_1$ satisfies the conditions of the arrow unification completeness lemma, Lemma 4.9, to get that $(S_1, \mathcal{A}_1) \vdash^\to \rho_1 = \sigma_0 \to \sigma_0' \Rightarrow (S_2, \mathcal{A}_2)$ such that $\exists R_2$ with $R_1 \cdot S_1 = R_2 \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$, $R_2 S_2 \sigma_0 = Q\sigma$ and $R_2 S_2 \sigma_0' = Q\sigma'$. Finally also $vars(R_2) \notin \mathcal{A}_2$ and $vars(S_2) \notin \mathcal{A}_2$ and similarly for the free variables of $\sigma_0 \to \sigma_0'$. Then also $S \cdot S_0 = R_2 \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$. Taking into account the freshness conditions as well, we can rewrite the premise of the rule $SS_0\Gamma \vdash_\Downarrow^{poly} u : Q\sigma$ as $R_2 S_2 \Gamma \vdash_\Downarrow^{poly} u : R_2 S_2 \sigma_0$. But now note that $\mathcal{A}_2$ is appropriate to apply the induction hypothesis to get that $(S_2, \mathcal{A}_2)\Gamma \vdash_\Downarrow^{poly} u : \sigma' \Rightarrow (S_3, \mathcal{A}_3)$, and $\exists R$ such that $R_2 \cdot S_2 = R \cdot S_3 \backslash_{\mathcal{A}_2 - \mathcal{A}_3}$, so $S \cdot S_0 = R \cdot S_3 \backslash_{\mathcal{A}_0 - \mathcal{A}_3}$. However we have that $\vdash^{inst} R_3 S_3 \sigma_0' \le \rho$ and $\mathcal{A}_3$ is appropriate to apply Lemma 4.18 to get that $S_3, \mathcal{A}_3)\Gamma \vdash^{inst} \sigma_0' \le \rho' \Rightarrow (S_3, \mathcal{A}_4)$ and exists R such that $R_3 \cdot S_3 = R \cdot S_4 \backslash_{\mathcal{A}_3 - \mathcal{A}_4}$ and $RS_4\rho' = \rho$. Moreover $vars(R) \notin \mathcal{A}_4$ and $vars(S_4) \notin \mathcal{A}_4$ and $ftv(\rho') \notin \mathcal{A}_4$. Therefore by applying the rule ALG-APP we are done, picking the same R.

- Case LET. In this case we have that $SS_0\Gamma \vdash_\Uparrow \mathtt{let}\ x = u\ \mathtt{in}\ t : \rho$ given that $SS_0\Gamma \vdash_\Uparrow^{poly} u : \sigma$ and $SS_0\Gamma, x : \sigma \vdash_\Uparrow t : \rho$. Consider an appropriate $\mathcal{A}_0$, such that $\mathcal{A}_0 \notin vars(S, S_0)$, $\mathcal{A}_0 \notin ftv(\Gamma)$, $\mathcal{A}_0 \notin ftv(\rho)$. Now we don't know that $\mathcal{A}_0 \notin ftv(\sigma)$, however let us consider the set of variables $\mathcal{X} = ftv(\sigma) - (ftv(\rho) \cup vars(S, S_0) \cup vars(\Gamma))$. Then consider a renaming substitution Q that take this set of variables to a disjoint from $\mathcal{A}_0$ set. Then by the substitution lemma we have that $SS_0\Gamma \vdash_\Uparrow^{poly} u : Q\sigma$ and has the same height. Similarly, applying the same Q to the other subderivation we get that $SS_0\Gamma, x : Q\sigma \vdash_\Uparrow t : \rho$ and has the same height. Now we are certain that the freshness conditions are met for the subderivations we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash^{poly} u : \sigma' \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R_1$ such that $S \cdot S_0 = R_1 \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$. Moreover $\vdash^{dsk} R_1 S_1 \sigma' \le_{sh} Q\sigma$ and $vars(R_1) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\sigma') \notin \mathcal{A}_1$. Then, by Lemma 3.19 we have that $SS_0\Gamma, x : R_1 S_1 \sigma' \vdash_\Uparrow t : \rho$ and this derivation has the same height as the derivation $SS_0\Gamma, x : Q\sigma \vdash_\Uparrow t : \rho$. Now, because $ftv(\Gamma) \notin \mathcal{A}_0 - \mathcal{A}_1$ it must be that $S \cdot S_0\Gamma = R_1 \cdot S_1\Gamma$. Then the derivation can be rewritten as $R_1 S_1\Gamma, x : R_1 S_1 \sigma' \vdash_\Uparrow t : \rho$. Then $\mathcal{A}_1$ is appropriate to apply the induction hypothesis and get that $(S_1, \mathcal{A}_1) ; \Gamma, x : \sigma' \vdash_\Uparrow t : \rho' \Rightarrow (S_2, \mathcal{A}_2)$ and $\exists R$ such that $R_1 \cdot S_1 = R \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$, and $RS_2\rho' = \rho$. Moreover $vars(R) \notin \mathcal{A}_3$, $vars(S_2) \notin \mathcal{A}_3$, $ftv(\rho') \notin \mathcal{A}_3$. Therefore $S \cdot S_0 = R \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$, and by the rule ALG-LET we get the result.

For the second part we have the following cases.

- Case VAR. The result follows again from Corollary 4.18.

- Case ABS2. Here we have $SS_0\Gamma \vdash_\Downarrow \backslash x . t : SS_0\rho'$, but since we are in the ABS2 case it must be that $SS_0\rho') = \sigma_a \to \sigma_r$ for some types σ_a and σ_r. Then consider a supply $\mathcal{A}_0$ that satisfies the freshness conditions. Therefore we have $\mathcal{A}_0 \notin vars(S, S_0)$ and $\mathcal{A}_0 \notin ftv(\rho') \cup ftv(\Gamma)$. Then it is easy to confirm

that the tape is appropriate for applying Lemma 4.9 to get that $(S_0, \mathcal{A}_0) \vdash^\rightarrow \rho' = \sigma_0 \rightarrow \sigma_0' \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R_u$ such that $S \cdot S_0 = R_u \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$ and $R_u S_1 \sigma_0 = \sigma_a$ and $R_u S_1 \sigma_0' = \sigma_r$. Moreover $vars(R_u) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\sigma_0) \notin \mathcal{A}_1$, $ftv(\sigma_0') \notin \mathcal{A}_1$. Then, because also $ftv(\Gamma) \notin \mathcal{A}_1$ the derivation can be rewritten as $R_u S_1 \Gamma \vdash_\Downarrow \backslash x \,.\, t : \sigma_a \rightarrow \sigma_r$ and the immediate subderivation was $R_u S_1 \Gamma, x : \sigma_a \vdash_\Downarrow^{poly} t : \sigma_r$. This can again be rewritten as $R_u S_1 \Gamma, x : R_u S_1 \sigma_0 \vdash_\Downarrow^{poly} t : R_u S_1 \sigma_0'$. Now, we know that $\mathcal{A}_1$ is appropriate to apply the induction hypothesis to get that $(S_1, \mathcal{A}_1) ; \Gamma, x : \sigma_0 \vdash_\Downarrow^{poly} t : \sigma_0' \Rightarrow (S_2, \mathcal{A}_2)$ and there exists an R, such that $R_u \cdot S_1 = R \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$ which implies $S \cdot S_0 = R \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$. Moreover $vars(R) \notin \mathcal{A}_2$, $vars(S_2) \notin \mathcal{A}_2$. Then we can apply the rule ALG-ABS2 to get the result and taking the same R finishes the case.

- Case AABS2. In this case we have that $SS_0 \Gamma \vdash_\Downarrow \backslash x :: \sigma_x \,.\, t : SS_0 \rho_0$, where $SS_0 \rho_0 = \sigma_a' \rightarrow \sigma_r'$ and $\vdash^{dsk} \sigma_a' \leq \sigma_x$ and $SS_0 \Gamma, x : \sigma_x \vdash_\Downarrow^{poly} t : \sigma_r'$. Consider a supply $\mathcal{A}_0$ that satisfies the freshness conditions, that is, $\mathcal{A}_0 \notin vars(S, S_0)$, $\mathcal{A}_0 \notin ftv(\Gamma) \cup ftv(\rho_0)$. Then we can verify that $\mathcal{A}_0$ is appropriate for applying Lemma 4.9 to get that $(S_0, \mathcal{A}_0) \vdash^\rightarrow \rho_0 = \sigma_a \rightarrow \sigma_r \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R_u$ such that $S \cdot S_0 = R_u \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, $R_u S_1 \sigma_a = \sigma_a'$, $R_u S_1 \sigma_r = \sigma_r'$, $vars(R_u) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\sigma_a) \notin \mathcal{A}_1$, $ftv(\sigma_r) \notin \mathcal{A}_1$. Then, we can rewrite $\vdash^{dsk} \sigma_a' \leq \sigma_x$ as $\vdash^{dsk} R_u S_1 \sigma_a \leq R_u S_1 \sigma_x$ because σ_x is closed. Then it is again not hard to verify that the tape $\mathcal{A}_1$ is appropriate to apply Lemma 4.17 to get that $(S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_a \leq \sigma_x \Rightarrow (S_2, \mathcal{A}_2)$ and $\exists R_1$ such that $R_u \cdot S_1 = R_1 \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$ which implies $S \cdot S_0 = R_1 \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$, $vars(R_1) \notin \mathcal{A}_2$ and $vars(S_1) \notin \mathcal{A}_2$. Then, from the premises of the rule we get that $R_1 S_2 \Gamma, x : \sigma_x \vdash_\Downarrow t : R_1 S_2 \sigma_r$. Now it is easy to verify that $\mathcal{A}_2$ in an appropriate supply to apply the induction hypothesis to get that $\mathcal{A}_2, S_2) \Gamma, x : \sigma_x \vdash_\Downarrow t : \sigma_r \Rightarrow (S_3, \mathcal{A}_3)$ and $\exists R$ with $R_1 S_2 = R \cdot S_3 \backslash_{\mathcal{A}_2 - \mathcal{A}_3}$ which implies $S \cdot S_0 = Q = \backslash_{\mathcal{A}_0 - \mathcal{A}_3}$; moreover $vars(R) \notin \mathcal{A}_3$ and $vars(S_3) \notin \mathcal{A}_3$. Then we have all the premises of rule ALG-AABS2 and by applying it we get the result; picking the same R finishes the case.

- Case APP. In this case we have $SS_0 \Gamma \vdash_\Uparrow t \, u : SS_0 \rho$, given that $SS_0 \Gamma \vdash_\Uparrow t : \sigma \rightarrow \sigma'$, $SS_0 \Gamma \vdash_\Downarrow^{poly} u : \sigma$ and $SS_0 \Gamma \vdash_\Downarrow^{inst} \sigma' \leq SS_0 \rho$. Consider any symbol supply $\mathcal{A}_0$ that satisfies the freshness conditions, that is, $\mathcal{A}_0 \notin vars(S, S_0)$, $\mathcal{A}_0 \notin ftv(\rho)$, $\mathcal{A}_0 \notin ftv(\Gamma)$. Then let us consider the set of variables $\mathcal{X} = ftv(\sigma \rightarrow \sigma') - (vars(S, S_0) \cup ftv(\rho) \cup ftv(\Gamma)$ and let Q be a renaming substitution taking $\mathcal{X}$ to set of variables disjoint from $\mathcal{A}_0$. Then by the substitution lemma we can create derivations: $SS_0 \Gamma \vdash_\Uparrow t : Q\sigma \rightarrow Q\sigma'$, $SS_0 \Gamma \vdash_\Downarrow^{poly} u : Q\sigma$ and $SS_0 \Gamma \vdash_\Downarrow^{inst} Q\sigma' \leq SS_0 \rho$ and these derivations have the same height as the original derivations. Now we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash_\Uparrow t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1)$ such that $\exists R_1$ with $S \cdot S_0 = R_1 \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$, $vars(R_1) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\rho_1) \notin \mathcal{A}_1$. $R_1 S_1 \rho_1 = Q\sigma \rightarrow Q\sigma'$. Now it is easy to verify that $\mathcal{A}_1$ satisfies the conditions of the unification completeness lemma, Lemma 4.9 to get that $(S_1, \mathcal{A}_1) \vdash^\rightarrow \rho_1 = \sigma_0 \rightarrow \sigma_0' \Rightarrow (S_2, \mathcal{A}_2)$ such that $\exists R_2$ with $R_1 \cdot S_1 = R_2 \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$, $R_2 S_2 \sigma_0 = Q\sigma$ and $R_2 S_2 \sigma_0' = Q\sigma'$. Finally also $vars(R_2) \notin \mathcal{A}_2$ and $vars(S_2) \notin \mathcal{A}_2$ and similarly for the free variables of $\sigma_0 \rightarrow \sigma_0$. Then also $S \cdot S_0 = R_2 \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$. Then, taking into account the freshness conditions as well, we can rewrite the premise of the rule $SS_0 \Gamma \vdash_\Downarrow^{poly} u : Q\sigma$ as $R_2 S_2 \Gamma \vdash_\Downarrow^{poly} u : R_2 S_2 \sigma_0$. But note that $\mathcal{A}_2$ is now appropriate to apply the induction hypothesis to get that $(S_2, \mathcal{A}_2) \Gamma \vdash_\Downarrow^{poly} u : \sigma_0 \Rightarrow (S_3, \mathcal{A}_3)$, and $\exists R_3$ such that $R_2 \cdot S_2 = R_3 \cdot S_3 \backslash_{\mathcal{A}_2 - \mathcal{A}_3}$, so $S \cdot S_0 = R_3 \cdot S_3 \backslash_{\mathcal{A}_0 - \mathcal{A}_3}$. Then we have that $\vdash_\Downarrow^{inst} Q\sigma' \leq SS_0 \rho$ and we can rewrite this as $\vdash^{dsk} R_3 S_3 \sigma_0' \leq R_3 S_3 \rho$ and now $\mathcal{A}_3$ is appropriate to apply Lemma 4.18 to get that $S_3, \mathcal{A}_3) \vdash_\Downarrow^{inst} \sigma_0' \leq \rho \Rightarrow (S_4, \mathcal{A}_4)$ such that $\exists R$ with $R_3 \cdot S_3 = R \cdot S_4 \backslash_{\mathcal{A}_3 - \mathcal{A}_4}$ which also gives us $S \cdot S_0 = R \cdot S_4 \backslash_{\mathcal{A}_0 - \mathcal{A}_4}$ and $vars(R) \notin \mathcal{A}_4$, $vars(S_4) \notin \mathcal{A}_4$. Then by applying the rule ALG-APP and taking the same R we are done.

- Case LET. In this case we have that $SS_0 \Gamma \vdash_\Downarrow \texttt{let } x \texttt{ = } u \texttt{ in } t : SS_0 \rho$ given that $SS_0 \Gamma \vdash_\Uparrow^{poly} u : \sigma$ and $SS_0 \Gamma, x : \sigma \vdash_\Downarrow t : SS_0 \rho$. Consider an appropriate $\mathcal{A}_0$, such that $\mathcal{A}_0 \notin vars(S, S_0)$, $\mathcal{A}_0 \notin ftv(\Gamma)$, $\mathcal{A}_0 \notin ftv(\rho)$. Now we don't know that $\mathcal{A}_0 \notin ftv(\sigma)$, however let us consider the set of variables $\mathcal{X} = ftv(\sigma) - (ftv(\rho) \cup vars(S, S_0) \cup vars(\Gamma))$. Then consider a renaming substitution Q that take this set of variables to a disjoint from $\mathcal{A}_0$ set. Then by the substitution lemma we have that $SS_0 \Gamma \vdash_\Uparrow^{poly} u : Q\sigma$ and has the same height. Similarly, applying the same Q to the other subderivation we get that

77

$SS_0\Gamma, x : Q\sigma \vdash_{\Downarrow} t : SS_0\rho$ and has the same height. Now we are certain that the freshness conditions are met for the subderivations we can apply the induction hypothesis to get that $(S_0, \mathcal{A}_0) ; \Gamma \vdash_{\Uparrow}^{poly} u : \sigma' \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R_1$ such that $S \cdot S_0 = R_1 \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$. Moreover $\vdash^{dsk} R_1 S_1 \sigma' \leq_{sh} Q\sigma$ and $vars(R_1) \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$, $ftv(\sigma') \notin \mathcal{A}_1$. Then, by Lemma 3.19 we have that $SS_0\Gamma, x : R_1 S_1 \sigma' \vdash_{\Downarrow} t : SS_0\rho$ and this derivation has the same height as the derivation $SS_0\Gamma, x : Q\sigma \vdash_{\Downarrow} t : SS_0\rho$. Now, because $ftv(\Gamma) \notin \mathcal{A}_0 - \mathcal{A}_1$ it must be that $SS_0\Gamma = R_1 S_1 \Gamma$ and similarly $SS_0\rho = R_1 S_1 \rho$. Then the derivation can be rewritten as $R_1 S_1 \Gamma, x : R_1 S_1 \sigma' \vdash_{\Downarrow} t : R_1 S_1 \rho$. Then $\mathcal{A}_1$ is appropriate to apply the induction hypothesis and get that $(S_1, \mathcal{A}_1) ; \Gamma, x : \sigma' \vdash_{\Downarrow} t : \rho \Rightarrow (S_2, \mathcal{A}_2)$ and $\exists R$ such that $R_1 \cdot S_1 = R \cdot S_2 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$. Moreover $vars(R) \notin \mathcal{A}_3$, $vars(S_2) \notin \mathcal{A}_3$. Therefore $S \cdot S_0 = R \cdot S_2 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$, and by the rule ALG-LET we get the result, taking the same R.

Let us consider the third part now. The rule used here was GEN1. We have that $SS_0\Gamma \vdash_{\Uparrow}^{poly} t : \forall \overline{b} . [\overline{a \mapsto b}]\rho$ given that $\Gamma \vdash_{\Uparrow}^{poly} t : \rho$, $\overline{a} = ftv(\rho) - ftv(\Gamma)$ and $\overline{b} \notin ftv(\rho) - \overline{a}$. The inferred polytype can be rewritten as $\sigma = \overline{SS_0\Gamma}(\rho)$. Pick an appropriate symbol supply $\mathcal{A}_0 \overline{b} \notin vars(S)$, $\mathcal{A}_0 \overline{b} \notin vars(S_0)$, $\mathcal{A}_0 \overline{b} \notin ftv(\rho) - (\overline{a} \cup \overline{\alpha})$, and $\mathcal{A}_0 \overline{b} \notin ftv(\Gamma)$. Then by induction hypothesis $\mathcal{A}_0, S_0) ; \Gamma \vdash_{\Uparrow} t : \rho' \Rightarrow (S_1, \mathcal{A}_1)$ and $\exists R$ such that $S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_1}$. Moreover $R S_1 \rho' = \rho$ and $ftv(\rho') \notin \mathcal{A}_1$, $vars(S_1) \notin \mathcal{A}_1$ and $vars(R) \notin \mathcal{A}_1$.

Now, by Lemma 4.15 we know that $fov(S_1 \rho') \subseteq fov(S_1 \Gamma) \subseteq fov(S_0 \Gamma)$. But what that says is that if $b \in fov(S_1 \rho')$ $b \in fov(S_0 \Gamma)$ and hence also $b \in fov(S_1 \Gamma)$. So **there cannot be any ordinary variables in the returned type that are not in the context**, this is why our notation uses $\overline{\alpha}$ instead of some general set symbol $\mathcal{X}$ or $\overline{\alpha} \cup \overline{a}$ in the ALG-GEN1 rule. This means that we can apply the rule ALG-GEN1 to get that $\mathcal{A}_0, S_0) ; \Gamma \vdash_{\Uparrow}^{poly} t : \sigma' \Rightarrow (S_1, \mathcal{A}_1)$ where $\sigma' = \overline{S_1 \Gamma}(S_1 \rho')$. We also pick the same R. Finally we need to show that $R S_1 \sigma' \leq_{sh} \sigma$. We have by Lemma 3.17 that

$$
\begin{aligned}
R S_1 \sigma' &= R S_1 \overline{S_1 \Gamma}(S_1 \rho') \\
&\leq_{sh} \overline{R S_1 S_1 \Gamma}(R S_1 S_1 \rho') \\
&= \overline{R S_1 \Gamma}(R S_1 \rho') \\
&= \overline{S S_0 \Gamma}(R S_1 \rho') = \overline{S S_0 \Gamma}(\rho) = \sigma
\end{aligned}
$$

For the fourth part the rule used was GEN2. It must be that $SS_0\Gamma \vdash_{\Downarrow}^{poly} t : SS_0\sigma$, given that

$$pr(SS_0\sigma) = \forall \overline{a} . \rho_a \tag{1}$$

$$\overline{a} \notin ftv(SS_0\Gamma) \tag{2}$$

$$SS_0\Gamma \vdash_{\Downarrow} t : \rho_a \tag{3}$$

Consider a supply $\mathcal{A}_0 \notin vars(S), vars(S_0), ftv(\Gamma, \sigma)$, By (1) and Lemma 4.16 we get

$$\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall \overline{b} . \rho_b \Rightarrow \mathcal{A}_1 \tag{4}$$

such that $SS_0(\forall \overline{b} . \rho_b) = \forall \overline{a} . \rho_a$. Moreover, by Lemma 4.11 we know that $\overline{b} \in \mathcal{A}_0$; hence $\overline{b} \notin vars(S, S_0)$. Then $SS_0\rho_b = [\overline{a \mapsto b}]\rho_a$. By the substitution lemma for (3) and taking into account (2) we get $SS_0\Gamma \vdash_{\Downarrow} t : SS_0\rho_b$. We can apply the induction hypothesis for the supply $\mathcal{A}_1$ since the last derivation has the same height as (3) to get that

$$(S_0, \mathcal{A}_1) ; \Gamma \vdash_{\Downarrow} t : \rho_b \Rightarrow (S_1, \mathcal{A}_2) \tag{5}$$

Morever there exists R such that $S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_1 - \mathcal{A}_2}$, or equivalently $S \cdot S_0 = R \cdot S_1 \backslash_{\mathcal{A}_0 - \mathcal{A}_2}$ and $vars(R) \notin \mathcal{A}_2$ and $vars(S_1) \notin \mathcal{A}_2$ as required. To apply rule ALG-GEN2 for (4) and (5) we only need to show that $\overline{b} \notin ftv(S_1 \Gamma, S_1 \sigma)$. Assume by contradiction that for some $b \in \overline{b}$ it is the case that $b \in ftv(S_1 \Gamma, S_1 \sigma)$. Then, since $b \notin ftv(\Gamma, \sigma)$ as $b \in \mathcal{A}_0$ it must be that there exists a $\gamma \in ftv(\Gamma, \sigma)$ such that $b \in S_1 \gamma$. Moreover $\gamma \notin \mathcal{A}_0$. Then $SS_0\gamma = RS_1\gamma$ and therefore $b \in vars(S, S_0)$, a contradiction to the freshness conditions of the supply. $\square$

4.6.2 Soundness

For this section only we are going to assume that meta variables can also occur in the types of the syntax-directed system; they are going to be treated as equivalent to ordinary variables.

Lemma 4.20 (Weak Prenex Conversion Soundness). *If $\mathcal{A}_0 \notin ftv(\sigma)$ and $\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall \overline{a}.\rho \Rightarrow \mathcal{A}_1$ then $pr(\sigma) = \forall \overline{a}.\rho$.*

Proof. Easy induction on the structure or σ. $\qquad\square$

Lemma 4.21 (Algorithmic Subsumption Soundness).

1. If $\mathcal{A}_0 \notin ftv(\sigma_1), \mathcal{A}_0 \notin ftv(\sigma_2)$, $\mathcal{A}_0 \notin vars(S_0)$, $(S_0, \mathcal{A}_0) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ then $\vdash^{dsk} S_1\sigma_1 \leq S_1\sigma_2$.

2. If $\mathcal{A}_0 \notin ftv(\sigma_1), \mathcal{A}_0 \notin ftv(\sigma_2)$, $\mathcal{A}_0 \notin vars(S_0)$, $(S_0, \mathcal{A}_0) \vdash^{dsk*} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$ then $\vdash^{dsk} S_1\sigma_1 \leq S_1\sigma_2$.

Proof. We prove the two claims simultaneously by induction the height of the derivations.

Part 1: We have the case of ASKOL to consider. We are given a supply $\mathcal{A}_0\overline{b}$ such that $\mathcal{A}_0\overline{b} \notin vars(S_0)$, $\mathcal{A}_0\overline{b} \notin ftv(\sigma_1, \sigma_2)$. We have that $(S_0, \mathcal{A}_0\overline{b}) \vdash^{dsk} \sigma_1 \leq \sigma_2 \Rightarrow (S_1, \mathcal{A}_2)$, given that

$$\mathcal{A}_0 \vdash^{pr} \sigma_2 \mapsto \forall \overline{a}.\rho \Rightarrow \mathcal{A}_1 \tag{1}$$

$$(S_0, \mathcal{A}_1) \vdash^{dsk*} \sigma_1 \leq \rho \Rightarrow (S_1, \mathcal{A}_2) \tag{2}$$

$$\overline{a} \notin ftv(S_1\sigma_1, S_1\sigma_2) \tag{3}$$

By (1) and Lemma (4.20) we get $pr(\sigma_2) = \forall \overline{a}.\rho$ and moreover we know by Lemma 4.11 that $\overline{a} \in \mathcal{A}_0$. Then, by induction hypothesis for (2) we get $\vdash^{dsk*} S_1\sigma_1 \leq S_1\rho$. it must be that $pr(S_1\sigma_2) = S_1(\forall \overline{a}.\rho$, and because of (3) $pr(S_1\sigma_2) = \forall \overline{a}.S_1\rho$. Moreover by (3) again we can apply rule SKOL to get that $\vdash^{dsk} S_1\sigma_1 \leq S_1\sigma_2$, as required.

Part 2: We have the following cases to consider.

- Case ASPEC. In this case we have that for some appropriate $\mathcal{A}_0\overline{\beta}$ that satisfies the freshness conditions $\mathcal{A}_0, S_0) \vdash^{dsk*} \forall \overline{a}.\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$ given that $\mathcal{A}_0, S_0) \vdash^{dsk*} [\overline{a} \mapsto \beta]\rho_1 \leq \rho_2 \Rightarrow (S_1, \mathcal{A}_1)$. By induction hypothesis we have that $\vdash^{dsk} S_1[\overline{a} \mapsto \beta]\rho_1 \leq S_1\rho_2$. Assume again that $\overline{a}$ are not in the free variables of the inputs of the judgement. Then this becomes $\vdash^{dsk} [\overline{a} \mapsto S_1\beta]S_1\rho_1 \leq S_1\rho_2$ and we can apply the rule SPEC to get the result, again noticing that we can commute S_1 and the quantifier.

- Case AFUN1. Here we have that $(S_0, \mathcal{A}_0) \vdash^{dsk*} \rho \leq \sigma_3 \rightarrow \sigma_4 \Rightarrow (S_3, \mathcal{A}_3)$ given that $(S_0, \mathcal{A}_0) \vdash^{\rightarrow} \rho = \sigma_1 \rightarrow \sigma_2 \Rightarrow (S_1, \mathcal{A}_1)$, $(S_1, \mathcal{A}_1) \vdash^{dsk} \sigma_3 \leq \sigma_1 \Rightarrow (S_2, \mathcal{A}_2)$, $(S_2, \mathcal{A}_2) \vdash^{dsk*} \sigma_2 \leq \sigma_4 \Rightarrow (S_3, \mathcal{A}_3)$. Now by the arrow unification soundness, Lemma 4.9 we have that $S_1\rho = S_1\sigma_1 \rightarrow S_1\sigma_2$ and moreover $\exists R_1$ such that $S_1 = R_1 \cdot S_0$ and $vars(S_1) \notin \mathcal{A}_1$. Then we can apply the induction hypothesis to get that $\vdash^{dsk} S_2\sigma_3 \leq S_2\sigma_1$ and moreover we know that $\exists R_2$ such that $S_2 = R_2 \cdot S_1$ and $vars(S_2) \notin \mathcal{A}_2$. Finally, again by induction we get that $\vdash^{dsk} S_3\sigma_2 \leq S_3\sigma_4$ and we know that $\exists R_3$ such that $S_3 = R_3 \cdot S_2$ and $vars(S_3) \notin \mathcal{A}_3$. Then we wish to show also that $\vdash^{dsk} S_3\rho \leq S_3\sigma_3 \rightarrow S_3\sigma_4$. Now we have that $S_3\rho = R_3R_2S_1\rho = R_3R_2S_1\sigma_1 \rightarrow R_3R_2S_1\sigma_2$. Then by the subsumption substitution lemma, Lemma 2.38, we get that $\vdash^{dsk} R_3S_2\sigma_3 \leq R_3S_2\sigma_1$ or $\vdash^{dsk} S_3\sigma_3 \leq R_3R_2S_1\sigma_1$ and we can apply the rule FUN to get the result.

- Case AFUN2. Similar to the case for AFUN1.

- Case MONO. In this case we have that $(S_0, \mathcal{A}_0) \vdash^{dsk*} \tau_1 \leq \tau_2 \Rightarrow (S_1, \mathcal{A}_0)$ given that $S_0 \vdash \tau_1 = \tau_2 \Rightarrow S_1$. By the unification completeness lemma, Lemma 4.8 we get that $S_1\tau_1 = S_1\tau_2$ and therefore $S_1\tau_1 = S_1\tau_2$; hence by rule mono $\vdash^{dsk} S_1\tau_1 \leq S_1\tau_2$.

$\square$

Lemma 4.22 (Algorithmic Soundness).

1. $\mathcal{A}_0 \notin ftv(\Gamma) \cup vars(S_0) \qquad\qquad \Rightarrow \quad S_1\Gamma \vdash_\Uparrow t : S_1\rho$
   $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
2. $\mathcal{A}_0 \notin ftv(\Gamma) \cup vars(S_0) \cup ftv(\rho) \quad \Rightarrow \quad S_1\Gamma \vdash_\Downarrow t : S_1\rho$
   $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Downarrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$
3. $\mathcal{A}_0 \notin ftv(\Gamma) \cup vars(S_0) \qquad\qquad \Rightarrow \quad S_1\Gamma \vdash_\Uparrow^{poly} t : S_1\sigma$
   $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$
4. $\mathcal{A}_0 \notin ftv(\Gamma) \cup vars(S_0) \cup ftv(\sigma) \quad \Rightarrow \quad S_1\Gamma \vdash_\Downarrow^{poly} t : S_1\sigma$
   $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Downarrow^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_1)$

Proof. By induction on the algorithmic relations. We proceed by case analysis on the last rule used.

For the first part we have to consider the following cases.

- Case ALG-VAR. Here we have that $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow x : \rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $x : \sigma \in \Gamma$, $(S_0, \mathcal{A}_0) \vdash_\Uparrow^{inst} \sigma \leq \rho \Rightarrow (S_1, \mathcal{A}_1)$, and by a simple inversion we can verify that $S_1 = S_0$ and, if $\sigma = \forall\overline{a}.\rho'$, it must be that $\rho = [\overline{a \mapsto \beta}]\rho'$ where $\overline{\beta} = \mathcal{A}_0 - \mathcal{A}_1$. Then $\vdash_\Uparrow^{inst} \sigma \leq \rho$ by rules INST and INST1, and by the substitution lemma $\vdash_\Uparrow^{inst} S_1\sigma \leq S_1\rho$. Then we can apply the rule VAR to get that $S_1\Gamma \vdash_\Uparrow x : S_1\rho$.

- Case ALG-ABS1. Here $(S_0, \mathcal{A}_0\beta);\Gamma; \vdash_\Uparrow \backslash x.t : \beta \to \rho \Rightarrow (S_1, \mathcal{A}_1)$, given that $(S_0, \mathcal{A}_0);\Gamma, x : \beta \vdash_\Uparrow t : \rho \Rightarrow (S_1, \mathcal{A}_1)$. By induction hypothesis we get that $S_1\Gamma, x : S_1 b \vdash_\Uparrow t : S_1\rho$ and $\exists R$ such that $S_1 = R \cdot S_0$, $vars(S_1) \subseteq ftv(\Gamma) \cup vars(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$ and $ftv(\rho) \subseteq ftv(\Gamma) \cup vars(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$. Then by applying the rule ABS we get that $S_1\Gamma \vdash_\Uparrow \backslash x.t : S_1\beta \to S_1\rho$ as required.

- Case ALG-AABS1. Similar to the ALG-ABS1 case.

- Case ALG-APP. Here we have that $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow t\ u : \rho \Rightarrow (S_4, \mathcal{A}_4)$, given that $(S_0, \mathcal{A}_0)\,;\Gamma \vdash_\Uparrow t : \rho_1 \Rightarrow (S_1, \mathcal{A}_1)$, $(S_1, \mathcal{A}_1) \vdash^\to \rho_1 = \sigma \to \sigma' \Rightarrow (S_2, \mathcal{A}_2)$, $(S_2, \mathcal{A}_2)\,;\Gamma \vdash_\Downarrow^{poly} u : \sigma \Rightarrow (S_3, \mathcal{A}_3)$, $(S_3, \mathcal{A}_3) \vdash_\Uparrow^{inst} \sigma' \leq \rho \Rightarrow (S_4, \mathcal{A}_4)$. By induction hypothesis we get that $S_1\Gamma \vdash_\Uparrow t : S_1\rho_1$ and moreover $\exists R_1$ such that $S_1 = R_1 \cdot S_1$ and $vars(S_1) \subseteq ftv(\Gamma) \cup vars(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$, $ftv(\rho_1) \subseteq ftv(\Gamma) \cup vars(S_0) \cup (\mathcal{A}_0 - \mathcal{A}_1)$. Then we have the conditions to apply the arrow unification soundness to get that $S_2\rho_1 = S_2\sigma \to S_2\sigma'$ and $\exists R_2$ such that $S_2 = R_2 \cdot S_1$ and $vars(S_2) \subseteq ftv(\rho_1) \cup vars(S_1) \cup (\mathcal{A}_1 - \mathcal{A}_2)$. Then the conditions are appropriate to apply the induction hypothesis to get that $S_3\Gamma \vdash_\Downarrow^{poly} u : S_3\sigma$ and $\exists R_3$ such that $S_3 = R_3 \cdot S_2$ and $vars(S_3) \subseteq ftv(\Gamma) \cup vars(S_2) \cup (\mathcal{A}_2 - \mathcal{A}_3) \cup ftv(\sigma)$. Finally note that $S_4 = S_3$, $\mathcal{A}_3 = \mathcal{A}_4\overline{\beta}$ and asssume that $\sigma' = \forall\overline{a}.\rho_0$. Then $\rho = [\overline{a \mapsto \beta}]\rho_0$ because the instantiation inference relation is the identity. Therefore also $\vdash_\Uparrow^{inst} S_4\sigma' \leq S_4\rho$. Now by applying the substitution lemma we have that $R_3 R_2 S_1\Gamma \vdash_\Uparrow t : R_3 R_2 S_1\rho_1$ or $S_4\Gamma \vdash_\Uparrow t : R_3(S_2\sigma' \to S_2\rho')$ or $S_4\Gamma \vdash_\Uparrow t : R_4\sigma' \to R_4\rho'$. Then we also have that $S_4\Gamma \vdash_\Downarrow^{poly} u : R_4\sigma'$ and we can apply the rule APP to get the result.

- Case ALG-LET. The case uses a similar argument as the ALG-APP case.

The second part is similar to the proof of the first part appealing to the algorithmic subsumption soundness lemma.

For the third part we have that $(S_0, \mathcal{A}_0\overline{b}) ; \Gamma \vdash_{\Uparrow}^{poly} t : \forall \overline{b}.[\overline{\alpha} \mapsto \overline{b}]S_1\rho \Rightarrow (S_1, \mathcal{A}_1)$ given that $\overline{\alpha} = ftv(S_1\rho) - ftv(S_1\Gamma)$ and $(S_0, \mathcal{A}_0);\Gamma \vdash_{\Uparrow} t : \rho \Rightarrow (S_1, \mathcal{A}_1)$. By induction hypothesis we get that $S_1\Gamma \vdash_{\Uparrow} t : S_1\rho$ and we know that for some R $S_1 = R \cdot S_1$, $ftv(\rho) \subseteq vars(S_0) \cup ftv(\Gamma) \cup (\mathcal{A}_0 - \mathcal{A}_1)$ and $vars(S_1) \subseteq vars(S_0) \cup ftv(\Gamma) \cup (\mathcal{A}_0 - \mathcal{A}_1)$. Then, we are certain that $\overline{b} \notin ftv(S_1\rho) - (\overline{\alpha})$ and we can apply the rule GEN1 to get the result.

For the last part we have a supply $\mathcal{A}_0$ such that $\mathcal{A}_0 \notin ftv(\Gamma, \sigma) \cup vars(S_0)$. And $(S_0, \mathcal{A}_0) ; \Gamma \vdash_{\Downarrow}^{poly} t : \sigma \Rightarrow (S_1, \mathcal{A}_2)$, given that

$$\mathcal{A}_0 \vdash^{pr} \sigma \mapsto \forall \overline{a}.\rho \Rightarrow \mathcal{A}_1 \tag{1}$$

$$(S_0, \mathcal{A}_1) ; \Gamma \vdash_{\Downarrow} t : \rho \Rightarrow (S_1, \mathcal{A}_2) \tag{2}$$

$$\overline{a} \notin ftv(S_1\Gamma, S_1\sigma) \tag{3}$$

By Lemma 4.20 we get that $pr(\sigma) = \forall \overline{a}.\rho$. We can apply the induction hypothesis for (2) to get that $S_1\Gamma \vdash_{\Downarrow} S_1\rho$. Additionally $pr(S_1\sigma) = S_1(\forall \overline{a}.\rho)$, but $\overline{a} \in \mathcal{A}_0$, but we know that $\overline{a} not in ftv(S_1\sigma$ therefore we can commute the quantifier and the substitution to get that $pr(S_1\sigma) = \forall \overline{a}.S_1\rho$. Finally to be able to apply rule GEN2* it must be that $\overline{a} \notin ftv(S_1\Gamma)$ which we have by (3). $\qquad\square$

As a corollary, returning to the original syntax-directed system where we did not allow meta type variables in the returned types, we can apply a ground substitution V that will map all the meta-variables of the types and the context to any monotype to get the result.

4.7 Principal Types

Here is the familiar principal types property for inference mode.
Theorem 4.23 (Principal Types).

1. If $\vdash_{\Uparrow} t : \rho$ there exists a ρ' such that for all ρ with $\vdash_{\Uparrow} t : \rho$ it is the case that $\vdash_{\Uparrow} t : \rho'$ and there exists a subtitution R such that $\rho = R\rho'$.

2. If $\vdash_{\Uparrow}^{poly} t : \sigma$ there exists a σ' such that for all σ with $\vdash_{\Uparrow}^{poly} t : \sigma$ it is the case that $\vdash_{\Uparrow}^{poly} t : \sigma'$ and $\vdash^{sh} \sigma' \leq \sigma$.

Proof. The first part follows by the completeness and soundness theorems. The second part is derived by the first part by inversion of $\vdash_{\Uparrow}^{poly}$ and the definition of $\vdash^{sh}$. $\qquad\square$

For checking mode a corresponding property is not true. Consider for example the two checking judgements below:

$$\vdash_{\Downarrow} \backslash g.(g\ 3, g\ \text{True}) : (\forall a.a \rightarrow \text{Int}) \rightarrow (\text{Int}, \text{Int})$$
$$\vdash_{\Downarrow} \backslash g.(g\ 3, g\ \text{True}) : (\forall a.a \rightarrow a) \rightarrow (\text{Int}, \text{Bool})$$

Suppose that there was a most general ρ such that

$$\vdash_{\Downarrow} \backslash g.(g\ 3, g\ \text{True}) : \rho$$

and $\vdash^{dsk} \rho \leq (\forall a.a \rightarrow \text{Int}) \rightarrow (\text{Int}, \text{Int})$, and $\vdash^{dsk} \rho \leq (\forall a.a \rightarrow a) \rightarrow (\text{Int}, \text{Bool})$. Then it must be that $\rho = \sigma_1 \rightarrow \sigma_2$ such that

$$\vdash^{dsk} \forall a.a \rightarrow \text{Int} \leq \sigma_1$$
$$\vdash^{dsk} \forall a.a \rightarrow a \leq \sigma_1$$
$$\vdash^{dsk} \sigma_2 \leq (\text{Int}, \text{Int})$$
$$\vdash^{dsk} \sigma_2 \leq (\text{Int}, \text{Bool})$$

Assuming that $pr(\sigma_1) = \forall \overline{a}.\rho_1$ by inversion it must be that

$$\vdash^{dsk} \forall a . a \to \mathtt{Int} \le \rho_1$$
$$\vdash^{dsk} \forall a . a \to a \le \rho_1$$

By inversion it must be that

$$\vdash^{dsk} \tau_1 \to \mathtt{Int} \le \rho_1$$
$$\vdash^{dsk} \tau_2 \to \tau_2 \le \rho_1$$

Now it is easy to confirm that ρ_1 must be $\mathtt{Int} \to \mathtt{Int}$, therefore also $\sigma_1 = \mathtt{Int} \to \mathtt{Int}$. Consequently we would have to check that

$$\vdash_\Downarrow \backslash g . (g\ 3, g\ \mathtt{True}) : (\mathtt{Int} \to \mathtt{Int}) \to \sigma_2$$

But the above would fail as g is used polymorphically.

References

[1] Luis Damas and Robin Milner. Principal type-schemes for functional programs. In *Conference Record of the 9th Annual ACM Symposium on Principles of Programming Languages*, pages 207–12, New York, 1982. ACM Press.

[2] J. Roger Hindley. *Basic simple type theory*. Cambridge University Press, New York, NY, USA, 1997.

[3] R Milner. A theory of type polymorphism in programming. *JCSS*, 13(3), December 1978.

[4] John C. Mitchell. Polymorphic type inference and containment. *Inf. Comput.*, 76(2-3):211–249, 1988.

[5] M Odersky and K Läufer. Putting type annotations to work. In *23rd ACM Symposium on Principles of Programming Languages (POPL'96)*, pages 54–67. ACM, St Petersburg Beach, Florida, January 1996.

[6] Simon Peyton Jones, Dimitrios Vytiniotis, Stephanie Weirich, and Mark Shields. Practical type inference for arbitrary-rank types. Submitted to the Journal of Functional Programming, 2005.