MAX-PLANCK-INSTITUT FÜR INFORMATIK

Proof Normalization and
Subject Reduction
in Extensions of Fsub

Sergei Vorobyov

MPI–I–95–2–001 January 1995

Im Stadtwald
D 66123 Saarbrücken
Germany

Author's Address

Sergei Vorobyov (sv@mpi-sb.mpg.de),
Max–Planck–Institut für Informatik
Im Stadtwald
D-66123 Saarbrücken
Germany
(http://www.mpi-sb.mpg.de/guide/staff/sv/sv.html)

Publication Notes

Acknowledgements

I greatly benefited from discussions with Martín Abadi and Benjamin Pierce.

Abstract

System $F_{\leq}$, the second-order polymorphic typed λ-calculus with subtyping appeared to be undecidable because of the undecidability of its subtyping component. The discovery of decidable extensions of the $F_{\leq}$-subtyping relation put forward a challenging problem of incorporating these extensions into an $F_{\leq}$-like typing in a decidable and coherent manner. In this paper we describe a family of systems combining the standard $F_{\leq}$-typing rules with converging hierarchies of decidable extensions of the $F_{\leq}$-subtyping and give decidable criteria for successful *proof normalization* and *subject reduction*.

Proof Normalization and Subject Reduction in Extensions of $F_{\leq}$

Sergei Vorobyov

Max-Planck-Institut für Informatik
Im Stadtwald, D-66123, Saarbrücken, Germany
(e-mail: sv@mpi-sb.mpg.de, Phone: (49) 681-302-5391, Fax: (49) 681-302-5401)

January 16, 1995

Abstract

System $F_{\leq}$, the second-order polymorphic typed λ-calculus with subtyping [CW85, BL90, BTCCS91, CG92, CMMS94], appeared to be undecidable because of the undecidability of its subtyping component [Pie92]. The discovery of decidable extensions of the $F_{\leq}$-subtyping relation [Vor94a, Vor95] put forward a challenging problem of incorporating these extensions into an $F_{\leq}$-like typing in a decidable and coherent manner. In this paper we describe a family of systems combining the standard $F_{\leq}$-typing rules with converging hierarchies of decidable extensions of the $F_{\leq}$-subtyping and give decidable criteria for successful *proof normalization* and *subject reduction*.

1 Introduction

The well-known second-order polymorphic typed λ-calculus with subtyping, system $F_{\leq}$ [CW85, BL90, BTCCS91, CG92, CMMS94], extends Girard's system **F** by 1) an auxiliary subtyping relation $\Gamma \vdash \sigma \leq \tau$, 2) restricting type substitutions in type quantification and abstraction to subtypes of a given type, $(\Lambda\alpha \leq \rho\,.\,t) \;:\; (\forall\alpha \leq \rho\,.\,\tau)$, and 3) possibility to promote a type of a term to all supertypes by using the so-called (*Subsumption*) rule:

$$\frac{\Gamma \vdash t \;:\; \sigma \qquad\qquad \Gamma \vdash \sigma \leq \tau}{\Gamma \vdash t \;:\; \tau} \qquad (Sub)$$

The $F_{\leq}$ subtyping relation $\Gamma \vdash \sigma \leq \tau$ is generated by a simple collection of inference rules, with the most spectacular one, allowing for subtyping two boundedly quantified universal types:

$$\frac{\Gamma \vdash \tau_1 \leq \sigma_1 \qquad\qquad \Gamma\,,\,\alpha \leq \tau_1 \vdash \sigma_2 \leq \tau_2}{\Gamma \vdash (\forall\alpha \leq \sigma_1\,.\,\sigma_2) \leq (\forall\alpha \leq \tau_1\,.\,\tau_2)} \qquad (All)$$

In contrast to the simply-typed λ-calculus and system **F**, where the typing proofs are uniquely determined, a term in $F_{\leq}$ may have (infinitely) many different proofs and types. But $F_{\leq}$ possesses a nice *minimal type* property, guaranteeing the existence of a minimal type for every $F_{\leq}$-typeable term [CG92]. This minimal type is a subtype of every other type of the same term in the same context, and can be extracted from the uniquely determined *normal* typing proof of the term. Therefore, for the $F_{\leq}$-typechecking, it would have been sufficient to compute a minimal type of a term and to verify whether it subtypes to a given type.

However, the $F_{\leq}$ subtyping relation was proved undecidable by B.Pierce [Pie92]. As a consequence, the typing relation is also undecidable.

Weakening the $F_{\leq}$ subtyping relation, by restricting or modifying the above quantifier rule (*All*), does not give completely satisfactory subsystems of $F_{\leq}$. For example, the system $F_{\leq}^{\top}$ [CP94], obtained by replacing τ_1 by the largest type $\top$ in the right premise of (*All*), fails to possess the minimal type property (G.Ghelli). Despite the decidability of the $F_{\leq}^{\top}$ subtyping relation, the decidability of typing in $F_{\leq}^{\top}$ is still open. Replacing the same occurrence of τ_1 by σ_1 yields the system with nontransitive subtyping and open decidability problem for subtyping [CP94].

A possibility to circumvent the undecidability of $F_{\leq}$ by *reinforcement* (as opposed to weakening) was suggested in [Vor94a]. It turns out that the $F_{\leq}$ subtyping theory is not essentially undecidable [TMR53], possessing infinitely many different consistent decidable extensions. Every such extension is closed with respect to the $F_{\leq}$ subtyping rules and proves all $F_{\leq}$-provable subtypings (plus, being decidable, necessarily something else, but not all subtyping judgments). This result was obtained by interpreting the $F_{\leq}$ subtyping in M. Rabin's [Rab69] (weak) monadic second-order theories of successors **(W)SnS** [Vor94a, Vor95]. However, the direct **(W)SnS**-interpretations were not satisfactory for typing. First, they subtyped differently structured types. Second, each of them proved $\vdash\ \top \leq \top \to \top$ and $\vdash\ \top \leq (\forall\gamma \leq \top\,.\,\top)$. Therefore, assuming the standard $F_{\leq}$ typing rules, every term is typeable with the $\top$ type. The direct **(W)SnS**-interpretations were made more subtle, to avoid the above drawbacks, by combining them with the $F_{\leq}$-like subtyping rules, [Vor95]. The main result of [Vor95] can be interpreted as a

Refinement Theorem. *Given a (decidable) extension $T \supset F_{\leq}$ of the $F_{\leq}$ subtyping relation, closed with respect to the $F_{\leq}$-subtyping rules, one can construct a "more tight" (decidable) extension T' such that $F_{\leq} \subset T' \subset T$ and T' is also closed with respect to the $F_{\leq}$-subtyping rules.* □

As a boundary case, for $T = F_{\leq}$ the refinement gives $T' = F_{\leq}$ as the only fixpoint. The premise of the Refinement Theorem is satisfied, in particular (but not only), by the original **(W)SnS**-interpretations of $F_{\leq}$ [Vor94a, Vor95]. Iterating the Refinement Theorem one gets better and better decidable approximations to the undecidable $F_{\leq}$-subtyping, and the genuine $F_{\leq}$-subtyping as a limit. The properties of these hierarchies of converging decidable subtyping theories were investigated in depth in [Vor94b]. In particular, subtyping judgments distinguishing levels k and $k+1$ were constructed. Such judgments were then used to construct counterexamples to the *substitution property* (which is true for $F_{\leq}$):

$$\Gamma\ \vdash\ (\forall\gamma \leq \sigma_1\,.\,\sigma_2) \leq (\forall\gamma \leq \tau_1\,.\,\tau_2)\ \&\ \Gamma\ \vdash\ \rho \leq \tau_1\ \Rightarrow\ \Gamma\ \vdash\ \sigma_2[\rho/\gamma] \leq \tau_2[\rho/\gamma]. \tag{1}$$

Generally, the premises of (1) live in a *higher* hierarchy level than the conclusion. Sufficient conditions guaranteeing that they all belong to the *same* level were studied in [Vor94b]. This non-standard subtyping behavior has far-going consequences on the corresponding typing relation [Vor94c] and, as we show in this paper, makes typing proof normalization and the subject reduction property quite sophisticated.

Some results on combining the hierarchies of decidable extensions of the $F_{\leq}$-subtyping with the standard $F_{\leq}$ typing rules are given in [Vor94c]. Notably, it is demonstrated that 1) all the resulting systems extend $F_{\leq}$, typing every $F_{\leq}$-typeable term, 2) the normal typing relation is decidable in all extensions, 3) proofs that do not normalize correspond to provably $F_{\leq}$-untypeable terms.

In this paper we address the problem of *subject reduction* in extensions of $F_{\leq}$, i.e., preservation of types of well-typed terms by the 1st- and 2nd-order β-reduction (η-reduction poses no additional problems and will be considered elsewhere). Successful subject reduction is a consequence of the following three principal properties:

1. proof normalization, especially of type applications: a normal proof of $\Gamma\ \vdash\ t : (\forall\gamma \leq \sigma_1\,.\,\sigma_2)$ and a proof of $\Gamma\ \vdash\ \theta \leq \sigma_1$ should transform to a normal proof of $\Gamma\ \vdash\ t\,\{\theta\} : \sigma_2[\theta/\gamma]$.
2. Term-in-Typing-substitution: $\Gamma, x : \sigma, \Gamma'\ \vdash\ t : \tau$ and $\Gamma\ \vdash\ s : \sigma$ should imply $\Gamma, \Gamma'\ \vdash\ t[s/x] : \tau$.

3. Type-in-Typing-substitution: $\Gamma,\ \alpha \leq \sigma,\ \Gamma' \ \vdash\ t : \tau$ and $\Gamma \ \vdash\ \rho \leq \sigma$ should imply $\Gamma,\ \Gamma'[\rho/\alpha] \ \vdash\ t[\rho/\alpha] : \tau[\rho/\alpha]$.

All these properties hold for $F_{\leq}$. In extensions, as the proof normalization depends on the substitution property (1), they may fail. But we are able to give the decidable criteria guaranteeing that they hold. These criteria are tautologically true for $F_{\leq}$ (we thus get another proof of the subject reduction property for $F_{\leq}$), and, being decidable, could be added to the notion of well-typedness. As a result, we get decidable extensions of the $F_{\leq}$-typing satisfying the subject reduction property.

The paper is organized as follows. After Section 2 on preliminaries, we define in Section 3 hierarchies of decidable extensions of the $F_{\leq}$-subtyping, and the corresponding typing systems in Section 4. Proof normalization is addressed in Sections 5–7. Sections 8–13 are devoted to type and term substitutions. Subject reduction is discussed in Sections 14–16. In Section 17 we consider the example suggested by B. Pierce.

Martín Abadi and Benjamin Pierce constantly supplied me with counterexamples and encouragement.

2 Preliminaries

Definition 2.1 *The set of* boundedly quantified types *is defined by the grammar:*

$$\mathbb{T} \ \equiv_{df}\ \mathbb{V} \mid \top \mid \mathbb{T} \to \mathbb{T} \mid \forall \mathbb{V} \leq \mathbb{T}\,.\,\mathbb{T}$$

where: 1) $\mathbb{V}$ *is a set of* type variables *further denoted by Greek letters* α, β, γ; *2)* $\top$ *is the largest type; 3)* $\to$ *is the functional type constructor; 4)* $\forall\alpha \leq \rho\,.\,\tau$ *is a* boundedly quantified universal *type: a function assigning to each subtype* σ *of* ρ, $\sigma \leq \rho$, *the type* $\tau[\sigma/\alpha]$ *obtained from* τ *by substituting* σ *instead of free occurrences of* α. *In* $\forall\alpha \leq \rho\,.\,\tau$ *the bound* ρ does not contain α *free. Further, variable or compound types are denoted by Greek letters* σ, τ, ρ, θ, ϕ, ψ.

Definition 2.2 *The second-order* polymorphic terms *are defined by:*

$$T \ \equiv_{df}\ V \mid \lambda V\ :\ \mathbb{T}\,.\,T \mid T\,T \mid \Lambda\mathbb{V} \leq \mathbb{T}\,.\,T \mid T\,\{\mathbb{T}\}\ ,$$

corresponding respectively to object variables, functional and type abstraction and application.

Definition 2.3 *A* context Γ *is a finite sequence of subtyping and typing assumptions on variables.* $Dom(\Gamma)$ *is the set of variables appearing to the left of* $\leq$ *and* $:$ *in a context* Γ, $FV(e)$ *is the set of free variables in expression* e. *Contexts are defined inductively: 1)* $\emptyset$ *is a context. 2) If* Γ *is a context,* $x \in V$, $x \notin Dom(\Gamma)$, $\sigma \in \mathbb{T}$, *and* $FV(\sigma) \subseteq FV(\Gamma)$ *then* $\Gamma,\ x\ :\ \sigma$ *is a context. 3) If* Γ *is a context,* $\alpha \in \mathbb{V}$, $\alpha \notin Dom(\Gamma)$, $\sigma \in \mathbb{T}$, *and* $FV(\sigma) \subseteq FV(\Gamma)$ *then* $\Gamma,\ \alpha \leq \sigma$ *is a context. Define* $\Gamma(\alpha) = \sigma$ *and* $\Gamma(x) = \sigma$ *if* Γ *contains* $\alpha \leq \sigma$ *or* $x : \sigma$ *respectively. Define* $\Gamma^*(\alpha)$ *as* $\Gamma(\alpha)$ *if the latter is not a variable, and as* $\Gamma^*(\Gamma(\alpha))$ *otherwise*[1].

A subtyping judgment *is a figure of the form* $\Gamma \ \vdash\ \sigma\ \leq\ \tau$, *where* $FV(\sigma) \cup FV(\tau) \subseteq Dom(\Gamma)$.

A typing judgment *is a figure of the form* $\Gamma \ \vdash\ t\ :\ \tau$ *where* $FV(t) \cup FV(\tau) \subseteq Dom(\Gamma)$.

A $\mathcal{V}$-$\forall$-subtyping *is a judgment of the form* $\Gamma \ \vdash\ \alpha \leq (\forall\gamma \leq \sigma_1\,.\,\sigma_2)$.

3 Stratified Subtyping

The following is a slightly modified (by adding upper indices) Curien-Ghelli's algorithmic variant of $F_{\leq}$ (subtyping component) [CG92].

[1]Note that $\Gamma^*(\alpha)$ is always either $\top$, or a universal, or an arrow type

Definition 3.1 *A stratified subtype relation* $\Gamma \vdash^{(k)} \sigma \leq \tau$ *($k \in \mathbb{N}$) is defined by the rules:*

$$\Gamma \vdash^{(k)} \alpha \leq \alpha \quad (Refl) \qquad\qquad \Gamma \vdash^{(k)} \tau \leq \top \quad (Top)$$

$$\frac{\Gamma \vdash^{(k)} \tau_1 \leq \sigma_1 \qquad \Gamma \vdash^{(k)} \sigma_2 \leq \tau_2}{\Gamma \vdash^{(k)} \sigma_1 \to \sigma_2 \leq \tau_1 \to \tau_2} \quad (Arrow)$$

$$\frac{\Gamma \vdash^{(k)} \tau_1 \leq \sigma_1 \qquad \Gamma, \alpha \leq \tau_1 \vdash^{(k)} \sigma_2 \leq \tau_2}{\Gamma \vdash^{(k)} (\forall \alpha \leq \sigma_1 . \sigma_2) \leq (\forall \alpha \leq \tau_1 . \tau_2)} \quad (All)$$

$$\frac{\Gamma \vdash^{(k)} \Gamma^*(\alpha) \leq (\forall \beta \leq \rho . \tau)}{\Gamma \vdash^{(k+1)} \alpha \leq (\forall \beta \leq \rho . \tau)} \quad (AlgTrans\text{-}\forall)$$

$$\frac{\Gamma \vdash^{(k)} \Gamma(\alpha) \leq \tau}{\Gamma \vdash^{(k)} \alpha \leq \tau} \quad \{\tau \text{ is not a } \forall\text{-type}\} \quad (AlgTrans)$$

plus a set $\mathcal{H}$ of additional axioms of the form $\Gamma \vdash^{(0)} \alpha \leq (\forall \gamma \leq \sigma_1 . \sigma_2)$, called $\forall$-hypotheses. □

Note that two principal differences, as compared to $F_{\leq}$, are: 1) different levels, k and $k+1$, in the premise and the conclusion of ($AlgTrans$-$\forall$), and 2) the set of additional axioms $\mathcal{H}$. Respectively, two possibilities to turn the above system into $F_{\leq}$ are: either 1) letting $k = \omega$ and $\mathcal{H}$ empty, or 2) letting $\mathcal{H}$ equal the set of $F_{\leq}$-provable judgments of the form $\Gamma \vdash \alpha \leq (\forall \gamma \leq \sigma_1 . \sigma_2)$. Other choices of $\mathcal{H}$ may lead to converging hierarchies of decidable extensions of the $F_{\leq}$-subtyping:

Theorem 3.2 ([Vor94b]) *If the set $\mathcal{H}$ of $\forall$-axioms is decidable, closed with respect to the scheme: $\Gamma \vdash^{(0)} \alpha \leq \sigma \in \mathcal{H}$ & $\Gamma \vdash^{(0)} \sigma \leq \tau \Rightarrow \Gamma \vdash^{(0)} \alpha \leq \tau \in \mathcal{H}$, and contains every $F_{\leq}$-provable $\mathcal{V}$-$\forall$-subtyping, then:*

1. *for every k, the relation $\Gamma \vdash^{(k)} \sigma \leq \tau$ is decidable, transitive, and extends $F_{\leq}$;*
2. *every $\Gamma \vdash^{(k)}$ properly extends $\Gamma \vdash^{(k+1)}$; $\Gamma \vdash_{F_{\leq}}$ is the intersection of all $\Gamma \vdash^{(k)}$.* □

A trivial example is given by $\mathcal{H} = \{\text{all } \mathcal{V}\text{-}\forall\text{-subtypings}\}$. An infinite number of different and less obvious choices are suggested by **(W)SnS**-interpretations of $F_{\leq}$ [Vor94a, Vor95], each of them leads to a decidable converging hierarchy extending $F_{\leq}$. For every k there exist judgments *distinguishing* levels k and $k+1$, i.e., $\Gamma \vdash^{(k)} \sigma \leq \tau$, but not $\Gamma \vdash^{(k+1)} \sigma \leq \tau$. One can also construct closed types σ, τ with the spectacular properties, impossible in $F_{\leq}$, (see [Vor94c]:

$$\gamma \leq \sigma \vdash^{(k)} \gamma \leq \tau, \tag{2}$$

$$\not\vdash^{(k)} \sigma \leq \tau. \tag{3}$$

The types σ, τ provide counterexamples to the following *substitution property*, true for $F_{\leq}$:

$$\Gamma \vdash (\forall \gamma \leq \sigma_1 . \sigma_2) \leq (\forall \gamma \leq \tau_1 . \tau_2) \;\&\; \Gamma \vdash \rho \leq \rho_1 \Rightarrow \Gamma \vdash \sigma_2[\rho/\gamma] \leq \tau_2[\rho/\gamma]$$

Consider, e.g., $\vdash^{(k)} (\forall \gamma \leq \top . \gamma) \leq (\forall \gamma \leq \sigma . \tau)$, which is equivalent, by (All) and (Top), to (2), hence true. But substituting $[\sigma/\gamma]$ in it yields (3).

In general, type substitution leads to the *level decreasing*. In the example above we have:

$$\vdash^{(k+1)} (\forall \gamma \leq \top . \gamma) \leq (\forall \gamma \leq \sigma . \tau) \Rightarrow \vdash^{(k)} \sigma \leq \tau \tag{4}$$

Such non-standard behavior of substitution make typing proof normalization and subject reduction quite subtle, subject to certain preconditions, as will be discussed in the sequel.

4 Normal Typing

Definition 4.1 *For every $k \in \mathbb{N}$ the* minimal normal $\vdash^{(k)}_{mn}$ *and* normal $\vdash^{(k)}_{n}$ typing relations *are defined by the following inference rules:*

$$\Gamma \ \vdash^{(k)}_{mn} \ x \ : \ \Gamma(x) \quad (\mu\text{-}Var)$$

$$\frac{\Gamma,\, \alpha \le \sigma \ \vdash^{(k)}_{mn} \ r \ : \ \tau}{\Gamma \ \vdash^{(k)}_{mn} \ (\Lambda\alpha \le \sigma \,.\, r) \ : \ (\forall\alpha \le \sigma \,.\, \tau)}\,(\mu\text{-}TAbs) \qquad \frac{\Gamma,\, x : \sigma \ \vdash^{(k)}_{mn} \ r \ : \ \tau}{\Gamma \ \vdash^{(k)}_{mn} \ (\lambda x : \sigma \,.\, r) \ : \ \sigma \to \tau}\,(\mu\text{-}FAbs)$$

$$\frac{\left\{\begin{array}{ll} \textit{either} & \Gamma \ \vdash^{(k)}_{mn} \ f \ : \ \sigma \to \tau \\ \textit{or} & \Gamma \ \vdash^{(k)}_{mn} \ f \ : \ \alpha, \quad \Gamma^*(\alpha) \equiv \sigma \to \tau \end{array}\right. \qquad \Gamma \ \vdash^{(k)}_{n} \ t \ : \ \sigma}{\Gamma \ \vdash^{(k)}_{mn} \ f\ t \ : \ \tau}\quad (\mu\text{-}FApp)$$

$$\frac{\left\{\begin{array}{ll} \textit{either} & \Gamma \ \vdash^{(k)}_{mn} \ r \ : \ (\forall\gamma \le \sigma_1 \,.\, \sigma_2) \\ \textit{or} & \Gamma \ \vdash^{(k)}_{mn} \ r \ : \ \alpha, \quad \Gamma^*(\alpha) \equiv (\forall\gamma \le \sigma_1 \,.\, \sigma_2) \end{array}\right. \qquad \Gamma \ \vdash^{(k)} \ \phi \le \sigma_1}{\Gamma \ \vdash^{(k)}_{mn} \ r\{\phi\} \ : \ \sigma_2[\phi/\gamma]}\quad (\mu\text{-}TApp)$$

$$\frac{\Gamma \ \vdash^{(k)}_{mn} \ r \ : \ \sigma \qquad \left\{\begin{array}{ll} \Gamma \ \vdash^{(k+1)} & \sigma \le \tau \quad \textit{if it is a } \mathcal{V}\text{-}\forall\textit{-subtyping} \\ \Gamma \ \vdash^{(k)} & \sigma \le \tau \quad \textit{otherwise} \end{array}\right.}{\Gamma \ \vdash^{(k)}_{n} \ r \ : \ \tau}\quad (Sub)$$

These rules are inspired by [CMMS94]. They incorporate the standard $F_{\le}$-typing rules with the first major difference that our rules are stratified by levels. By forgetting upper indices one gets the rules of [CMMS94]. The second difference is the presence of two cases of $(Subsumption)$, with the distinguished case of $\mathcal{V}$-$\forall$-subtyping, corresponding to the subtyping rule $(AlgTrans\text{-}\forall)$. The case of $F_{\le}$ is covered by letting $\vdash^{(k)} \equiv \vdash_{F_{\le}}$-subtyping. Other possibilities are supplied by using hierarchies of subtyping relations described in Section 3.

Definition 4.2 *Define $\mu(\Gamma \ \vdash^{(k)} \ t)$, the* minimal type *of t in $\Gamma \ \vdash^{(k)}$, as the type σ such that $\Gamma \ \vdash^{(k)}_{mn} \ t : \sigma$ (if it exists). Define $\pi(\Gamma \ \vdash^{(k)} \ t)$, the* principal type *of t in $\Gamma \ \vdash^{(k)}$, as:*

$$\pi(\Gamma \ \vdash^{(k)} \ t) \ \equiv_{df} \ \left\{\begin{array}{ll} \mu(\Gamma \ \vdash^{(k)} \ t), & \textit{if } \mu(\Gamma \ \vdash^{(k)} \ t) \textit{ is not a variable,} \\ \Gamma^*(\mu(\Gamma \ \vdash^{(k)} \ t)), & \textit{if } \mu(\Gamma \ \vdash^{(k)} \ t) \textit{ is a variable,} \\ \textit{undefined,} & \textit{if } \mu(\Gamma \ \vdash^{(k)} \ t) \textit{ is undefined.} \end{array}\right.$$

Theorem 4.3 ([Vor94c]) *Relations $\Gamma \ \vdash^{(k)}_{mn} \ t \ : \ \theta$, $\Gamma \ \vdash^{(k)}_{n} \ t \ : \ \theta$ are decidable, minimal and principal types are computable iff the subtyping relation $\Gamma \ \vdash^{(k)} \ \sigma \le \tau$ is decidable.*

If τ is minimal for $\Gamma \ \vdash \ t$ in $F_{\le}$, then also for every $k \in \mathbb{N}$ one has $\mu(\Gamma \ \vdash^{(k)} \ t) = \tau$. □

5 Normalization

Typing proof normalization is the transformation of non-normal typing proofs into normal ones. Normalization can be described by the schemes like:

$$\frac{\Gamma \vdash_n^{(k)} f : \sigma \to \tau \qquad \Gamma \vdash_n^{(k)} t : \sigma}{\Gamma \vdash_n^{(k)} f\,t : \tau} \quad (\mathcal{N}\text{-}FApp)$$

which means that given two normal proofs $\Gamma \vdash_n^{(k)} f : \sigma \to \tau$ and $\Gamma \vdash_n^{(k)} t : \sigma$ one can *always* transform them into a normal proof of $\Gamma \vdash_n^{(k)} f\,t : \tau$. Note that the above scheme is *absent* from the set of rules for normal typing (stronger left premise $\Gamma \vdash_{mn}^{(k)} f : \sigma \to \tau$ is stipulated in (μ-*FApp*)). The schemes as ($\mathcal{N}$-*FApp*), provided admissible, can be seen as meta-theorems extending the set of rules for normal typing, or as derivable rules for normal typing.

Given a non-normal proof, one can normalize it, using the normalization schemes. At each step it suffices to choose maximal subproofs that are already normal, and to apply a corresponding normalization scheme, if available. A very useful by-product of typing proof normalization in $F_{\leq}$ is that it automatically yields the minimal types for terms. Given a normal typing proof, it suffices to retrieve the type *cut* by the *last* subsumption applied in this proof. This type is a subtype of every other type of the same term in the same context, see [CG92].

In the extensions of $F_{\leq}$, the normalization of all constructs, *except type application*, is easy, see Lemma 5.1. Normalization of type applications is quite subtle and may fail, in general, see Section 6. A criterion for the successful normalization of type applications is given by Lemma 7.1

Lemma 5.1 (Normalization) *The following are valid normalization schemes:* ($\mathcal{N}$-*FApp*)*,*

$$\frac{\Gamma, x:\sigma \vdash_n^{(k)} t : \tau}{\Gamma \vdash_n^{(k)} (\lambda x:\sigma\,.\,t) : \sigma \to \tau}\,(\mathcal{N}\text{-}FAbs) \qquad \frac{\Gamma, \alpha \leq \sigma \vdash_n^{(k)} t : \tau}{\Gamma \vdash_n^{(k)} (\Lambda\alpha \leq \sigma\,.\,t) : (\forall\alpha \leq \sigma\,.\,\tau)}\,(\mathcal{N}\text{-}TAbs)$$

Proof. ($\mathcal{N}$-*FApp*): The normal proof of $\Gamma \vdash_n^{(k)} f : \sigma \to \tau$ is either

$$\frac{\Gamma \vdash_{mn}^{(k)} f : \alpha \qquad \Gamma \vdash^{(k)} \alpha \leq \sigma \to \tau}{\Gamma \vdash_n^{(k)} f : \sigma \to \tau} \quad \text{or} \quad \frac{\Gamma \vdash_{mn}^{(k)} f : \sigma' \to \tau' \qquad \Gamma \vdash^{(k)} \sigma' \to \tau' \leq \sigma \to \tau}{\Gamma \vdash_n^{(k)} f : \sigma \to \tau}$$

In the first case, let $\sigma' \to \tau' \equiv \Gamma^*(\alpha)$. From $\Gamma \vdash^{(k)} \alpha \leq \sigma \to \tau$ by (*AlgTrans*) we have: $\Gamma \vdash^{(k)} \sigma' \to \tau' \leq \sigma \to \tau$ (the same as in the second case). Hence, in both cases, by (*Arrow*):

$$\Gamma \vdash^{(k)} \sigma \leq \sigma', \tag{5}$$
$$\Gamma \vdash^{(k)} \tau' \leq \tau. \tag{6}$$

Consequently, the normal proof of $\Gamma \vdash_n^{(k)} f\,t : \tau$ may be constructed as follows:

$$\frac{\dfrac{\begin{cases}\text{either } \Gamma \vdash_{mn}^{(k)} f : \sigma' \to \tau', \\ \text{or } \Gamma \vdash_{mn}^{(k)} f : \alpha,\ \Gamma^*(\alpha) \equiv \sigma' \to \tau'\end{cases} \qquad \dfrac{\Gamma \vdash_n^{(k)} t : \sigma \qquad \overset{See\ (5)}{\Gamma \vdash^{(k)} \sigma \leq \sigma'}}{\Gamma \vdash_n^{(k)} t : \sigma'}}{\Gamma \vdash_{mn}^{(k)} f\,t : \tau'} \qquad \overset{See\ (6)}{\Gamma \vdash^{(k)} \tau' \leq \tau}}{\Gamma \vdash_n^{(k)} f\,t : \tau}$$

($\mathcal{N}$-*TAbs*): The normal proof of $\Gamma, \alpha \leq \sigma \vdash_n^{(k)} t : \tau$ is:

$$\frac{\Gamma, \alpha \leq \sigma \vdash_{mn}^{(k)} t : \tau' \qquad \begin{cases}\text{either} & \Gamma, \alpha \leq \sigma \vdash^{(k+1)} \tau' \overset{(*J1*)}{\leq} \tau, & \text{if it is a } \mathcal{V}\text{-}\forall\text{-subtyping} \\ \text{or} & \Gamma, \alpha \leq \sigma \vdash^{(k)} \tau' \leq \tau, & \text{if it is not a } \mathcal{V}\text{-}\forall\text{-subtyping}\end{cases}}{\Gamma, \alpha \leq \sigma \vdash_n^{(k)} t : \tau}$$

Since $(*J1*)$ always implies, by (All), $\Gamma \vdash^{(k)} (\forall\alpha \le \sigma\,.\,\tau') \le (\forall\alpha \le \sigma\,.\,\tau)$, the normal proof of $\Gamma \vdash_n^{(k)} (\Lambda\alpha \le \sigma\,.\,t) : (\forall\alpha \le \sigma\,.\,\tau)$ is as follows:

$$\dfrac{\dfrac{\Gamma,\, \alpha \le \sigma \vdash_{mn}^{(k)} t : \tau'}{\Gamma,\, \alpha \le \sigma \vdash_{mn}^{(k)} (\Lambda\alpha \le \sigma\,.\,t) : (\forall\alpha \le \sigma\,.\,\tau')} \qquad \Gamma \vdash^{(k)} (\forall\alpha \le \sigma\,.\,\tau') \le (\forall\alpha \le \sigma\,.\,\tau)}{\Gamma,\, \alpha \le \sigma \vdash_n^{(k)} (\Lambda\alpha \le \sigma\,.\,t) : (\forall\alpha \le \sigma\,.\,\tau)}$$

$(\mathcal{N}\text{-}FAbs)$: The proof is completely analogous to the preceding one. □

6 Failure of the Type Application Normalization

One of the normal typing rules for the type application is:

$$\dfrac{\Gamma \vdash_{mn}^{(k)} r : (\forall\gamma \le \sigma_1\,.\,\sigma_2) \qquad \Gamma \vdash^{(k)} \theta \le \sigma_1}{\Gamma \vdash_{mn}^{(k)} r\,\{\theta\} : \sigma_2[\theta/\gamma]} \qquad (\mu\text{-}TApp)$$

The corresponding rule for functional application $(\mu\text{-}FApp)$ gives rise to the normalization scheme $(\mathcal{N}\text{-}FApp)$ from Section 5, which differs from the rule just by replacement of $\Gamma \vdash_{mn}^{(k)}$ with $\Gamma \vdash_n^{(k)}$ in the left premise and conclusion of $(\mu\text{-}FApp)$. The similarly obtained normalization scheme for $(\mu\text{-}TApp)$:

$$\dfrac{\Gamma \vdash_{n}^{(k)} r : (\forall\gamma \le \sigma_1\,.\,\sigma_2) \qquad \Gamma \vdash^{(k)} \theta \le \sigma_1}{\Gamma \vdash_{n}^{(k)} r\,\{\theta\} : \sigma_2[\theta/\gamma]} \qquad (\mathcal{N}\text{-}TApp) \tag{7}$$

is *not valid*, in general, in extensions of $F_{\le}$ (but is valid for the $F_{\le}$ itself!). Consider a simple counterexample. Let σ, τ be the closed types with properties (2), (3) from Section 3. Then

$$r : (\forall\gamma \le \top\,.\,\gamma) \vdash_n^{(k)} r : (\forall\gamma \le \sigma\,.\,\tau) \tag{8}$$

is provable, since the necessary subtyping

$$\vdash^{(k)} (\forall\gamma \le \top\,.\,\gamma) \le (\forall\gamma \le \sigma\,.\,\tau) \tag{9}$$

is equivalent, by (All), to (2). But applying the scheme (7) to (8) and $(Refl)$:

$$\dfrac{r : (\forall\gamma \le \top\,.\,\gamma) \vdash_n^{(k)} r : (\forall\gamma \le \sigma\,.\,\tau) \qquad\qquad r : (\forall\gamma \le \top\,.\,\gamma) \vdash^{(k)} \sigma \le \sigma}{r : (\forall\gamma \le \top\,.\,\gamma) \vdash_n^{(k)} r\,\{\sigma\} : \tau}$$

yields the unprovable judgment, since, by $(\mu\text{-}TApp)$, $r : (\forall\gamma \le \top\,.\,\gamma) \vdash_{mn}^{(k)} r\,\{\sigma\} : \sigma$, and the necessary subtyping $r : (\forall\gamma \le \top\,.\,\gamma) \vdash^{(k)} \sigma \le \tau$ is unprovable by (3).

The existence of pairs of types with the properties (2), (3) is immanent for hierarchical extensions of the $F_{\le}$-subtyping [Vor94b, Vor94c]. Therefore, the general failure of $(\mathcal{N}\text{-}TApp)$ is common for all hierarchical extensions of $F_{\le}$.

7 Successful Normalization of Type Applications

The following lemma establishes the criterion for the successful normalization of type applications in the hierarchical extensions of $F_{\le}$. Note that the condition (10) is *tautologically true* for $F_{\le}$. Thus type applications normalize unconditionally in $F_{\le}$, which together with Lemma 5.1 gives the proof of the proof normalization theorem for $F_{\le}$ [CG92]. Note also that the condition (10) is effectively decidable in all decidable extensions of Section 3. Finally, if a term is $F_{\le}$-typeable, its typing proof is normalizable in every extension. Conversely, if it is non-normalizable in some extension, the term is guaranteed to be $F_{\le}$-untypeable.

Lemma 7.1 (Type Application Normalization Criterion)

$$\frac{\Gamma \vdash_n^{(k)} r \ : \ (\forall\gamma \le \sigma_1 \,.\, \sigma_2) \overset{(*J1*)}{} \qquad \Gamma \vdash^{(k)} \theta \le \sigma_1 \overset{(*J2*)}{}}{\Gamma \vdash_n^{(k)} r\,\{\theta\} \ : \ \sigma_2[\theta/\gamma] \overset{(*J3*)}{}} \quad (\mathcal{N}\text{-}TApp)$$

if and only if the principal type $(\forall\gamma \le \rho_1 \,.\, \rho_2) \ \equiv \ \pi(\Gamma \vdash^{(k)} r)$ *satisfies the condition:*

$$\begin{cases} \Gamma \vdash^{(k+1)} \ \rho_2[\theta/\gamma] \ \le \ \sigma_2[\theta/\gamma], & \textit{if it is a } \mathcal{V}\textit{-}\forall\textit{-subtyping,} \\ \Gamma \vdash^{(k)} \ \rho_2[\theta/\gamma] \ \le \ \sigma_2[\theta/\gamma], & \textit{if it is not a } \mathcal{V}\textit{-}\forall\textit{-subtyping.} \end{cases} \tag{10}$$

Proof. $(*J1*)$ by $(\mu\text{-}TApp)$ and (Sub) implies $\Gamma \vdash^{(k)} (\forall\gamma \le \rho_1 \,.\, \rho_2) \ \le \ (\forall\gamma \le \sigma_1 \,.\, \sigma_2)$, hence, by (All), $\Gamma \vdash^{(k)} \sigma_1 \le \rho_1$, which together with $(*J2*)$ gives $\Gamma \vdash^{(k)} \theta \le \rho_1$. (11)

Therefore, the normal proof of $(*J3*)$ looks like:

$$\frac{\dfrac{\pi(\Gamma \vdash^{(k)} r) \equiv (\forall\gamma \le \rho_1 \,.\, \rho_2) \qquad \overset{See\ (11)}{\Gamma \vdash^{(k)} \theta \le \rho_1}}{\Gamma \vdash_{mn}^{(k)} r\,\{\theta\} \ : \ \rho_2[\theta/\gamma]} \qquad \{\ (10)\ \}}{\Gamma \vdash_n^{(k)} r\,\{\theta\} \ : \ \sigma_2[\theta/\gamma]} \quad (Sub)$$

Returning to the counterexample of Section 6, it is interesting to see where it violates the conditions of the Lemma. The principal type of r in the counterexample is: $\forall\gamma \le \rho_1 \,.\, \rho_2 \equiv (\forall\gamma \le \top \,.\, \gamma)$, the type $(\forall\gamma \le \sigma_1 \,.\, \sigma_2) \equiv (\forall\gamma \le \sigma \,.\, \tau)$, and $\theta \equiv \sigma$. The Lemma stipulates, see (10), that $\Gamma \vdash^{(k)} \rho_2[\theta/\gamma] \equiv \sigma \ \le \ \tau \equiv \sigma_2[\theta/\gamma]$. But this is false, see (3), by the choice of σ and τ.

8 Term-in-Typing and Type-in-Typing Substitutions

Term-in-Typing-substitution property, i.e., possibility to derive

$$\frac{\Gamma, x : \sigma, \Gamma' \vdash t \ : \ \tau \qquad \Gamma \vdash s \ : \ \sigma}{\Gamma, \Gamma' \vdash t[s/x] \ : \ \tau} \tag{12}$$

and *Type-in-Typing-substitution* property, i.e., possibility to derive

$$\frac{\Gamma, \alpha \le \sigma, \Gamma' \vdash t \ : \ \tau \qquad \Gamma \vdash \rho \le \sigma}{\Gamma, \Gamma'[\rho/\alpha] \vdash t\,[\rho/\alpha] \ : \ \tau\,[\rho/\alpha]} \tag{13}$$

are fundamental for typed λ-calculi and indispensable for proving the subject reduction property.

Term-in-Typing- and Type-in-Typing-substitution properties are non-obvious when typing proofs are *not uniquely determined*, e.g., in the presence of subtyping. Directly substituting one normal proof into another may yield non-normal proofs. Therefore, substitution should be accompanied by a *proof normalization*. As we saw in Sections 6, 7, successful normalization in extensions of $F_{\le}$ should be subject to certain preconditions. Consequently, some preconditions have to be stipulated to guarantee the correctness of the Term-in-Typing- and Type-in-Typing-substitutions. The problem becomes more sophisticated when the explicit analysis of *typing levels* is involved.

9 Substitution Properties May Fail in Extensions of $F_{\leq}$

Let σ, τ, be types defined in Section 3 and used to construct the counterexample of Section 6 with the properties (2), (3). Consider the valid judgments:

$$g : (\forall\gamma \leq \top\,.\,\gamma),\, f : (\forall\gamma \leq \sigma\,.\,\tau) \quad \vdash_{mn}^{(k)} \quad f\,\{\sigma\} \;:\; \tau \tag{14}$$

$$g : (\forall\gamma \leq \top\,.\,\gamma) \quad \vdash_{mn}^{(k)} \quad g : (\forall\gamma \leq \top\,.\,\gamma) \tag{15}$$

From (15) and (9) we get, by (Sub), $g : (\forall\gamma \leq \top\,.\,\gamma) \;\vdash_{n}^{(k)}\; g : (\forall\gamma \leq \sigma\,.\,\tau)$.

Therefore, g is "*good*" for substituting instead of f in (14). But the resulting judgment

$$g : (\forall\gamma \leq \top\,.\,\gamma) \;\vdash_{n}^{(k)}\; (f\,\{\sigma\})\,[g/f] \equiv g\,\{\sigma\} \;:\; \tau$$

is *not* valid, since $g : (\forall\gamma \leq \top\,.\,\gamma) \;\vdash_{mn}^{(k)}\; g\,\{\sigma\} \;:\; \sigma$, but, by (3), $\not\vdash^{(k)}\; \sigma \leq \tau$. This counterexample shows that the Term-in-Typing-substitution (12) fails, in general in extensions of $F_{\leq}$.

Similarly, for the Type-in-Typing-substitution, always with the same types σ and τ, the next particular case of the scheme (13) has its premises true and the conclusion false:

$$\frac{\gamma \leq \sigma,\, f : (\forall\gamma \leq \top\,.\,\gamma) \;\vdash_{n}^{(k)}\; f\,\{\gamma\} \;:\; \tau \qquad\qquad \vdash^{(k)}\; \sigma \leq \sigma}{f : (\forall\gamma \leq \top\,.\,\gamma) \vdash_{n}^{(k)}\; f\,\{\sigma\} \;:\; \tau}$$

since $f : (\forall\gamma \leq \top\,.\,\gamma) \vdash_{mn}^{(k)} \; f\,\{\sigma\} \;:\; \sigma$, and for the conclusion of the last figure to be true one needs $\vdash^{(k)}\; \sigma \leq \tau$, which is false by (3). This gives the counterexample to Type-in-Typing-substitution (13).

10 Safety for Term-in-Typing Substitution

Definition 10.1 (Safety I) *Let* $\Gamma \;\vdash_{n}^{(k)}\; s \;:\; \sigma$. *A provable typing judgment*

$$\Gamma,\, x : \sigma \;\vdash_{n}^{(k)}\; t \;:\; \tau \tag{16}$$

is called $[s/x]$-safe *iff for every context term* $\Gamma,\, x : \sigma,\, \Gamma' \;\vdash^{(k)}\; r\,\{\theta\}$ *that occurs in the proof of* (16) *and for the principal types*

$$\begin{aligned} (\forall\gamma \leq \sigma_1\,.\,\sigma_2) \quad &\equiv_{df} \quad \pi(\Gamma,\, x : \sigma,\, \Gamma' \;\vdash^{(k)}\; r), \\ (\forall\gamma \leq \rho_1\,.\,\rho_2) \quad &\equiv_{df} \quad \pi(\Gamma,\, \Gamma' \;\vdash^{(k)}\; r[s/x]) \end{aligned}$$

the following condition is satisfied:

$$\begin{cases} \Gamma,\, \Gamma' \;\vdash^{(k+1)} & \rho_2[\theta/\gamma] \;\leq\; \sigma_2[\theta/\gamma], & \textit{if it is a } \mathcal{V}\textit{-}\forall\textit{-subtyping}, \\ \Gamma,\, \Gamma' \;\vdash^{(k)} & \rho_2[\theta/\gamma] \;\leq\; \sigma_2[\theta/\gamma], & \textit{if it is not a } \mathcal{V}\textit{-}\forall\textit{-subtyping}. \end{cases} \tag{17}$$

11 Term-in-Typing Substitution Lemma

Lemma 11.1 (Term-in-Typing Substitution) *Suppose that* $\Gamma \;\vdash_{n}^{(k)}\; s \;:\; \sigma$.

$$\textit{Then} \quad \Gamma,\, \Gamma' \;\vdash_{n}^{(*JX*)(k)}\; t[s/x] : \tau \quad \textit{iff} \quad \Gamma,\, x : \sigma,\, \Gamma' \;\vdash_{n}^{(*JY*)(k)}\; t : \tau \quad \textit{is} \quad [s/x]\textit{-safe}.$$

Proof. By induction on complexity of the term t. If t is a variable, then the conclusion is immediate. Proofs of necessity and of all inductive cases, except type application, are straightforward and omitted here.

Inductive Hypothesis. Suppose that the conclusion of the Lemma holds for all proper subterms of t.

Type Application. Let $t \equiv r\,\{\theta\}$. We have to construct the proof of $(*JX*)$. The proof of $(*JY*)$ is:

$$\frac{\Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(*J1*)}_{mn}{}^{(k)}\; r\,\{\theta\} \;:\; \sigma_2[\theta/\gamma] \qquad \left\{\begin{array}{lll} \Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k+1)} & \overset{(*J2*)}{\sigma_2[\theta/\gamma] \le \tau}, & \text{for } \mathcal{V}\text{-}\forall\text{-subtyping} \\ \Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)} & \sigma_2[\theta/\gamma] \le \tau, & \text{otherwise} \end{array}\right.}{\Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)}_{n}\; r\,\{\theta\} \;:\; \tau}$$

where $(\forall\gamma \le \sigma_1\,.\,\sigma_2)$ is the principal type of r in $\Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)}$. The proof of $(*J1*)$ is:

$$\frac{\left\{\begin{array}{lll} \text{either} & \Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)}_{mn} & \overset{(*J3*)}{r \;:\; \alpha \text{ and } (\Gamma,\,\Gamma')^*(\alpha) \equiv (\forall\gamma \le \sigma_1\,.\,\sigma_2)} \\ \text{or} & \Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)}_{mn} & r \;:\; (\forall\gamma \le \sigma_1\,.\,\sigma_2) \end{array}\right. \qquad \Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(*J4*)}{}^{(k)}\; \theta \le \sigma_1}{\Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(*J1*)}_{mn}{}^{(k)}\; r\,\{\theta\} \;:\; \sigma_2[\theta/\gamma]}$$

Anyway, both cases in $(*J3*)$ imply $\Gamma,\, x:\sigma,\, \Gamma' \;\vdash^{(k)}_{n}\; r \;:\; (\forall\gamma \le \sigma_1\,.\,\sigma_2)$. (18)

Applying the inductive hypothesis to (18) we get: $\Gamma,\, \Gamma' \;\vdash^{(k)}_{n}\; r[s/x] \;:\; (\forall\gamma \le \sigma_1\,.\,\sigma_2)$. (19)

The assumption $x:\sigma$ in the context of $(*J4*)$ can be omitted: $\Gamma,\, \Gamma' \;\vdash^{(k)}\; \theta \le \sigma_1$. (20)

From (19) and (20), by Lemma 7.1, since $(*JY*)$ is $[s/x]$-safe, we derive:

$$\Gamma,\, \Gamma' \;\vdash^{(k)}_{n}\; (r[s/x])\{\theta\} \equiv (r\{\theta\})[s/x] \;:\; \sigma_2[\theta/\gamma]. \tag{21}$$

(21) and $(*J2*)$ (with $x:\sigma$ omitted, subtyping does not depend on typing) yield $(*JX*)$. □

12 Safety for Type-in-Typing Substitution

Definition 12.1 (Safety II) *Let* $\Gamma \;\vdash^{(k)}\; \rho \le \sigma$. *A provable typing judgment*

$$\Gamma,\, \alpha \le \sigma \;\vdash^{(k)}_{n}\; t \;:\; \tau \tag{22}$$

is called $[\rho/\alpha]$-safe *iff for every context term* $\Gamma,\, \alpha \le \sigma,\, \Gamma' \;\vdash^{(k)}\; r\;\{\theta\}$ *that occurs in the proof of* (22) *and for the principal types*

$$\begin{array}{lll} (\forall\gamma \le \sigma_1\,.\,\sigma_2) & \equiv_{df} & \pi(\Gamma,\, \alpha \le \sigma,\, \Gamma' \;\vdash^{(k)}\; r), \\ (\forall\gamma \le \rho_1\,.\,\rho_2) & \equiv_{df} & \pi(\Gamma,\, \Gamma' \;\vdash^{(k)}\; r[\rho/\alpha]) \end{array}$$

the following condition is satisfied:

$$\left\{\begin{array}{lll} \Gamma,\, \Gamma'[\rho/\alpha] \;\vdash^{(k+1)} & \rho_2[\theta/\gamma] \;\le\; \sigma_2[\theta/\gamma], & \textit{if it is a } \mathcal{V}\textit{-}\forall\textit{-subtyping}, \\ \Gamma,\, \Gamma'[\rho/\alpha] \;\vdash^{(k)} & \rho_2[\theta/\gamma] \;\le\; \sigma_2[\theta/\gamma], & \textit{if it is not a } \mathcal{V}\textit{-}\forall\textit{-subtyping}. \end{array}\right. \tag{23}$$

and all subtyping judgments in the proof of (22) *are* $[\rho/\alpha]$-stable, *i.e.,*

$$\Gamma,\, \rho \le \alpha,\, \Gamma'[\rho/\alpha] \;\vdash^{(l)}\; \phi \le \psi \;\Rightarrow\; \Gamma,\, \Gamma'[\rho/\alpha] \;\vdash^{(l)}\; \phi[\rho/\alpha] \le \psi[\rho/\alpha] \tag{24}$$

Counterexample (4) of Section 3 shows that stability is important.

13 Type-in-Typing Substitution Lemma

Lemma 13.1 (Type-in-Typing Substitution) *Suppose that* $\Gamma \vdash^{(k)} \rho \leq \sigma$.

$$\textit{Then} \quad \Gamma, \Gamma'[\rho/\alpha] \vdash_n^{(k)} t[\rho/\alpha] : \tau[\rho/\alpha] \quad \textit{iff} \quad \Gamma, \alpha \leq \sigma, \Gamma' \vdash_n^{(k)} t : \tau \quad \textit{is } [\rho/\alpha]\textit{-safe.}$$

Proof. By induction on complexity of the term t. If t is a variable, the conclusion is immediate. Proofs of necessity and all inductive cases, except type application, are straightforward and skipped.

Inductive Hypothesis. Suppose that the conclusion of the Lemma holds for all proper subterms of t.

Type Application. Let $t \equiv r\,\{\theta\}$. We have to construct the proof of $(*JU*)$. The proof of $(*JV*)$ is:

$$\frac{\Gamma, \alpha \leq \sigma, \Gamma' \vdash_{mn}^{(k)} r\,\{\theta\} : \sigma_2[\theta/\gamma] \qquad \begin{cases} \Gamma, \alpha \leq \sigma, \Gamma' \vdash^{(k+1)} \sigma_2[\theta/\gamma] \leq \tau, & \text{for } \mathcal{V}\text{-}\forall\text{-subtyping} \\ \Gamma, \alpha \leq \sigma, \Gamma' \vdash^{(k)} \sigma_2[\theta/\gamma] \leq \tau, & \text{otherwise} \end{cases}}{\Gamma, \alpha \leq \sigma, \Gamma' \vdash_n^{(k)} r\,\{\theta\} : \tau}$$

where $(\forall\gamma \leq \sigma_1\,.\,\sigma_2)$ is the principal type of r in $\Gamma, \alpha \leq \sigma, \Gamma' \vdash^{(k)}$. The proof of $(*J1*)$ is:

$$\frac{\pi(\Gamma, \alpha \leq \sigma, \Gamma' \vdash^{(k)} r) \equiv (\forall\gamma \leq \sigma_1\,.\,\sigma_2) \qquad \Gamma, \alpha \leq \sigma, \Gamma' \vdash^{(k)} \theta \leq \sigma_1}{\Gamma, \alpha \leq \sigma, \Gamma' \vdash_{mn}^{(k)} r\,\{\theta\} : \sigma_2[\theta/\gamma]}$$

$(*J3*)$ implies: $\Gamma, \alpha \leq \sigma, \Gamma' \vdash_n^{(k)} r : (\forall\gamma \leq \sigma_1\,.\,\sigma_2)$.

Therefore, by inductive hypothesis, $\Gamma, \Gamma'[\rho/\alpha] \vdash_n^{(k)} r[\rho/\alpha] : (\forall\gamma \leq \sigma_1[\rho/\alpha]\,.\,\sigma_2[\rho/\alpha])$. (25)

From $(*J4*)$ by stability condition (24), $\Gamma, \Gamma'[\rho/\alpha] \vdash^{(k)} \theta[\rho/\alpha] \leq \sigma_1[\rho/\alpha]$. (26)

From (25) and (26), by Lemma 7.1, since $(*JV*)$ is $[\rho/\alpha]$-safe, we derive:

$$\Gamma, \Gamma'[\rho/\alpha] \vdash_n^{(k)} (r\,\{\theta\})[\rho/\alpha] : (\sigma_2[\theta/\gamma])[\rho/\alpha]. \tag{27}$$

It remains to apply the stability condition (24) to $(*J2*)$ to get:

$$\Gamma, \Gamma' \vdash^{(k+1)} (\sigma_2[\theta/\gamma])[\rho/\alpha] \leq \tau[\rho/\alpha] \quad \text{or} \quad \Gamma, \Gamma' \vdash^{(k)} (\sigma_2[\theta/\gamma])[\rho/\alpha] \leq \tau[\rho/\alpha] \tag{28}$$

The desired $(*JU*)$ is now derivable from (27) and (28) by (Sub). □

14 Subject Reduction Lemmas

are easy corollaries to Substitution Lemmas 11.1 and 13.1.

Lemma 14.1 ((β^1)-redex/reduct type preservation) *For arbitrary $k \in \mathbb{N}$ consider the uniquely determined normal typing proof of a (β^1)-redex $\Gamma \vdash_n^{(k)} (\lambda x : \sigma \,.\, t)\ s \ : \ \theta$:*

$$\dfrac{\dfrac{\dfrac{\overset{(*JA*)}{\Gamma,\, x : \sigma \vdash_{mn}^{(k)} t : \tau}}{\Gamma \vdash_{mn}^{(k)} (\lambda x : \sigma \,.\, t) \ : \ \sigma \to \tau} \qquad \Gamma \vdash_n^{(k)} s : \sigma}{\Gamma \vdash_{mn}^{(k)} (\lambda x : \sigma \,.\, t)\ s \ : \ \tau} \qquad \begin{cases} \Gamma \vdash^{(k+1)} \tau \leq \theta & \textit{for } \mathcal{V}\textit{-}\forall\textit{-subtyping} \\ \Gamma \vdash^{(k)} \tau \leq \theta & \textit{otherwise} \end{cases}}{\Gamma \vdash_n^{(k)} (\lambda x : \sigma \,.\, t)\ s \ : \ \theta}$$

One has the type preservation property for the (β^1)-reduction:

$$\Gamma \vdash_n^{(k)} (\lambda x : \sigma \,.\, t)\ s \ : \ \theta \ \Rightarrow \ \Gamma \vdash_n^{(k)} t[s/x] \ : \ \theta$$

if and only if the typing judgment $(*JA*)$ is $[s/x]$-safe. □

Definition 14.2 ((β^1) Redex Safety) *In notations of Lemma 14.1, call a (β^1) redex* $\Gamma \vdash_n^{(k)} (\lambda x : \sigma \,.\, t)\ s \ : \ \theta$ safe *iff the corresponding typing judgment $(*JA*)$ is $[s/x]$-safe.* □

Lemma 14.3 ((β^2)-redex/reduct type preservation) *For arbitrary $k \in \mathbb{N}$ consider the uniquely determined normal typing proof of a (β^2)-redex $\Gamma \vdash_n^{(k)} (\Lambda\alpha \leq \sigma \,.\, t)\,\{\rho\} \ : \ \theta$:*

$$\dfrac{\dfrac{\dfrac{\overset{(*JB*)}{\Gamma,\, \alpha \leq \sigma \vdash_{mn}^{(k)} t : \tau}}{\Gamma \vdash_{mn} (\Lambda\alpha \leq \sigma \,.\, t) : (\forall\alpha \leq \sigma \,.\, \tau)} \quad \Gamma \vdash^{(k)} \rho \leq \sigma}{\Gamma \vdash_{mn}^{(k)} (\Lambda\alpha \leq \sigma \,.\, t)\,\{\rho\} : \tau[\rho/\alpha]} \qquad \begin{cases} \Gamma \vdash^{(k+1)} \tau[\rho/\alpha] \leq \theta & \textit{for } \mathcal{V}\textit{-}\forall\textit{-subtyping} \\ \Gamma \vdash^{(k)} \tau[\rho/\alpha] \leq \theta & \textit{otherwise} \end{cases}}{\Gamma \vdash_n^{(k)} (\Lambda\alpha \leq \sigma \,.\, t)\,\{\rho\} \ : \ \theta}$$

One has the type preservation property for the (β^2)-reduction:

$$\Gamma \vdash_n^{(k)} (\Lambda\alpha \leq \sigma \,.\, t)\,\{\rho\} \ : \ \theta \ \Rightarrow \ \Gamma \vdash_n^{(k)} t[\rho/\alpha] \ : \ \theta$$

if and only if the typing judgment $(*JB*)$ is $[\rho/\alpha]$-safe. □

Definition 14.4 ((β^2) Redex Safety) *In notations of Lemma 14.3, call a (β^2) redex* $\Gamma \vdash_n^{(k)} (\Lambda\alpha \leq \sigma \,.\, t)\,\{\rho\} \ : \ \theta$ safe *iff the corresponding typing judgment $(*JB*)$ is $[\rho/\alpha]$-safe.* □

The following is immediate.

Proposition 14.5 *If the underlying subtyping theory is decidable, then the safety of (β^1) and (β^2) redices is decidable.* □

15 Reduction Relation

We now combine the Reduction Lemmas 14.1 and 14.3 to obtain the Subject Reduction property in its conventional form (as was asked by Martín Abadi), i.e.,

$$\Gamma \vdash t : \tau \;\; \textit{and} \;\; t \; \textit{reduces to} \; t' \;\; \textit{imply} \;\; \Gamma \vdash t' : \tau$$

First, some effort is needed to define the reduction relation, which is generated by the usual 1st- and 2nd-order beta-reductions and is additionally closed with respect to usual congruences (subterm reduction) and transitivity (many-step reduction).

Definition 15.1 (Reduction Relation) *For every $k \in \mathbb{N}$ the* reduction relation $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ *is defined by the following rules (2 axioms: (β^1), (β^2) and 6 rules: $(Trans)$, $(\lambda\text{-}Cong)$, $(FApp\text{-}Cong\text{-}L)$, $(FApp\text{-}Cong\text{-}R)$, $(\Lambda\text{-}Cong)$, $(TApp\text{-}Cong)$):*

$$\frac{\Gamma \vdash_n^{(k)} (\lambda x : \sigma \,.\, t)\, s : \theta \quad is\ safe}{\Gamma \vdash^{(k)} (\lambda x : \sigma \,.\, t)\, s \rightsquigarrow t[s/x] : \theta} \qquad (\beta^1)$$

$$\frac{\Gamma \vdash_n^{(k)} (\Lambda\alpha \le \sigma \,.\, t)\, \{\rho\} : \theta \quad is\ safe}{\Gamma \vdash^{(k)} (\Lambda\alpha \le \sigma \,.\, t)\, \{\rho\} \rightsquigarrow t[\rho/\alpha] : \theta} \qquad (\beta^2)$$

$$\frac{\Gamma \vdash^{(k)} r \rightsquigarrow s : \tau \qquad \Gamma \vdash^{(k)} s \rightsquigarrow t : \tau}{\Gamma \vdash^{(k)} r \rightsquigarrow t : \tau} \qquad (Trans)$$

$$\frac{\Gamma, x : \sigma \vdash^{(k)} t \rightsquigarrow t' : \tau}{\Gamma \vdash^{(k)} (\lambda x : \sigma \,.\, t) \rightsquigarrow (\lambda x : \sigma \,.\, t') : \sigma \to \tau} \qquad (\lambda\text{-}Cong)$$

$$\frac{\Gamma \vdash^{(k)} t \rightsquigarrow t' : \sigma \to \tau \qquad \Gamma \vdash_n^{(k)} s : \sigma}{\Gamma \vdash^{(k)} t\, s \rightsquigarrow t'\, s : \tau} \qquad (FApp\text{-}Cong\text{-}L)$$

$$\frac{\Gamma \vdash_n^{(k)} t : \sigma \to \tau \qquad \Gamma \vdash^{(k)} s \rightsquigarrow s' : \sigma}{\Gamma \vdash^{(k)} t\, s \rightsquigarrow t\, s' : \tau} \qquad (FApp\text{-}Cong\text{-}R)$$

$$\frac{\Gamma, \alpha \le \sigma \vdash^{(k)} t \rightsquigarrow t' : \tau}{\Gamma \vdash^{(k)} (\Lambda\alpha \le \sigma \,.\, t) \rightsquigarrow (\Lambda\alpha \le \sigma \,.\, t') : (\forall\alpha \le \sigma \,.\, \tau)} \qquad (\Lambda\text{-}Cong)$$

$$\frac{\Gamma \vdash^{(k)} r^1 \rightsquigarrow r^2 : (\forall\alpha \le \sigma \,.\, \tau) \qquad \Gamma \vdash^{(k)} \rho \le \sigma}{\Gamma \vdash^{(k)} r^1\, \{\rho\} \rightsquigarrow r^2\, \{\rho\} : \tau[\rho/\alpha]} \qquad (TApp\text{-}Cong)$$

$$\left[\begin{array}{l} \textit{provided that for } i = 1, 2,\ (\forall\alpha \le \rho_1^i \,.\, \rho_2^i) \equiv \pi(\Gamma \vdash^{(k)} r^i) \textit{ one has:} \\ \left\{\begin{array}{ll} \Gamma \vdash^{(k+1)} & \rho_2^i[\rho/\alpha] \le \tau[\rho/\alpha], \quad \textit{if it is a } \mathcal{V}\text{-}\forall\textit{-subtyping,} \\ \Gamma \vdash^{(k)} & \rho_2^i[\rho/\alpha] \le \tau[\rho/\alpha], \quad \textit{if it is not a } \mathcal{V}\text{-}\forall\textit{-subtyping.} \end{array}\right. \end{array}\right] \qquad (29)$$

A judgment of the form $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ is called a reduction judgment. □

The only premises of the rules (β^1) and (β^2) are typing judgments with additional safety conditions (decidable if the underlying subtyping theory $\Gamma \vdash^{(k)} \sigma \leq \tau$ is decidable). The right premise of $(FApp\text{-}Cong\text{-}L)$ and the left premise of $(FApp\text{-}Cong\text{-}R)$ are also typing judgments. The right premise of $(TApp\text{-}Cong)$ is a subtyping judgment. All other premises and conclusions are reduction judgments. The congruence rule $(TApp\text{-}Cong)$ is subject to the precondition (29) (also decidable if the underlying subtyping theory $\Gamma \vdash^{(k)} \sigma \leq \tau$ is decidable). As we mentioned above, the safety preconditions in (β^1), (β^2), and the precondition (29) of $(TApp\text{-}Cong)$ are tautologically true for $F_{\leq}$. Therefore, the above reduction system generates the usual (β^1)-(β^2)-reduction relation of $F_{\leq}$ if the underlying subtyping theory is one of $F_{\leq}$[2].

We will extensively use the following auxiliary

Lemma 15.2 *If* $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ *then* $\Gamma \vdash_n^{(k)} t : \theta$.

Proof. By induction on the structure of the inference of the reduction judgment $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ in the system from Definition 15.1. Every proof of a reduction judgment starts from axioms (β^1), (β^2) and consists of applications of the rules $(Trans)$, $(\lambda\text{-}Cong)$, $(FApp\text{-}Cong\text{-}L)$, $(FApp\text{-}Cong\text{-}R)$, $(\Lambda\text{-}Cong)$, $(TApp\text{-}Cong)$.

Basis. If $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ by (β^1) or (β^2), then the needed $\Gamma \vdash_n^{(k)} t : \theta$ is guaranteed by the premise of (β^1) or (β^2) respectively.

Inductive Hypothesis. Suppose the claim of the Lemma holds for all reduction judgments appearing in the proof of $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ (except itself!) in the system of Definition 15.1.

To prove that the claim holds also for $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ itself, we make the case analysis on a last rule from Definition 15.1 applied in the proof of $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$.

Case of $(Trans)$**:** Let the judgment $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ be obtained from $\Gamma \vdash^{(k)} t \rightsquigarrow t'' : \theta$ (A) and $\Gamma \vdash^{(k)} t'' \rightsquigarrow t' : \theta$ by $(Trans)$. From (A), by inductive hypothesis, $\Gamma \vdash_n^{(k)} t : \theta$, which is needed.

Case of $(\lambda\text{-}Cong)$**:** Let the reduction judgment

$$\Gamma \vdash^{(k)} t \equiv (\lambda x : \sigma \,.\, r) \rightsquigarrow (\lambda x : \sigma \,.\, r') \equiv t' : \sigma \to \tau \equiv \theta$$

be obtained from $\Gamma, x : \sigma \vdash^{(k)} r \rightsquigarrow r' : \tau$ (A) by $(\lambda\text{-}Cong)$. From (A), by inductive hypothesis, $\Gamma, x : \sigma \vdash_n^{(k)} r : \tau$ (B). Therefore, from (B), by Normalization Lemma 5.1,
$\Gamma \vdash_n^{(k)} t \equiv (\lambda x : \sigma \,.\, r) : \sigma \to \tau \equiv \theta$, which is needed.

Case of $(FApp\text{-}Cong\text{-}L)$**:** Let $\Gamma \vdash^{(k)} t \equiv r\,s \rightsquigarrow r'\,s \equiv t' : \tau$ be derived from $\Gamma \vdash^{(k)} r \rightsquigarrow r' : \sigma \to \tau$ (A) and $\Gamma \vdash_n^{(k)} s : \sigma$ (B) by $(FApp\text{-}Cong\text{-}L)$. From (A), by inductive hypothesis, $\Gamma \vdash_n^{(k)} r : \sigma \to \tau$ (C). Therefore, (C) and (B), by Normalization Lemma 5.1, yield the desired $\Gamma \vdash_n^{(k)} t \equiv r\,s : \tau$.

Case of $(FApp\text{-}Cong\text{-}R)$**:** Let $\Gamma \vdash^{(k)} t \equiv r\,s \rightsquigarrow r\,s' \equiv t' : \tau$ be derived from $\Gamma \vdash_n^{(k)} r : \sigma \to \tau$ (A) and $\Gamma \vdash^{(k)} s \rightsquigarrow s' : \sigma$ (B) by $(FApp\text{-}Cong\text{-}R)$. From (B), by inductive hypothesis, $\Gamma \vdash_n^{(k)} s : \sigma$ (C). Therefore, (A) and (C), by Normalization Lemma 5.1, yield the needed $\Gamma \vdash_n^{(k)} t \equiv r\,s : \tau$.

Case of $(\Lambda\text{-}Cong)$**:** Let the reduction judgment

$$\Gamma \vdash^{(k)} t \equiv (\Lambda\alpha \leq \sigma \,.\, r) \rightsquigarrow (\Lambda\alpha \leq \sigma \,.\, r') \equiv t' : (\forall\alpha \leq \sigma \,.\, \tau) \equiv \theta$$

be obtained from $\Gamma, \alpha \leq \sigma \vdash^{(k)} r \rightsquigarrow r' : \tau$ (A) by $(\Lambda\text{-}Cong)$. From (A), by inductive hypothesis, $\Gamma, \alpha \leq \sigma \vdash_n^{(k)} r : \tau$ (B). Therefore, from (B), by Normalization Lemma 5.1,
$\Gamma \vdash_n^{(k)} t \equiv (\Lambda\alpha \leq \sigma \,.\, r) : (\forall\alpha \leq \sigma \,.\, \tau) \equiv \theta$, which is needed.

[2] Cf., Paul Dirac's correspondence principle.

Case of $(TApp\text{-}Cong)$**:** Let $\Gamma \vdash^{(k)} t \equiv r\,\{\rho\} \rightsquigarrow r'\,\{\rho\} \equiv t' : \tau[\rho/\alpha]$ be derived from $\Gamma \vdash^{(k)} r \rightsquigarrow r' : (\forall\alpha \le \sigma\,.\,\tau)$ (A) and $\Gamma \vdash^{(k)} \rho \le \sigma$ (B) by $(TApp\text{-}Cong)$. From (A), by inductive hypothesis, $\Gamma \vdash_n^{(k)} r : (\forall\alpha \le \sigma\,.\,\tau)$ (C). Therefore, (C) and (B), by Normalization Lemma 7.1 (because of (29)) yield the desired $\Gamma \vdash_n^{(k)} t \equiv r\,\{\rho\} : \tau[\rho/\alpha]$. □

16 Subject Reduction Theorem

Theorem 16.1 (Subject Reduction) *Suppose* $\Gamma \vdash_n^{(k)} t : \theta$ *and* $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$. *Then* $\Gamma \vdash_n^{(k)} t' : \theta$. □

Remark. By Lemma 15.2, the first assumption of the Theorem may be skipped, but we keep it for the sake of standardness. □

Proof. By induction on the structure of the inference of the reduction judgment $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ in the system from Definition 15.1. Every proof of a reduction judgment starts from axioms (β^1), (β^2) and consists of applications of the rules $(Trans)$, $(\lambda\text{-}Cong)$, $(FApp\text{-}Cong\text{-}L)$, $(FApp\text{-}Cong\text{-}R)$, $(\Lambda\text{-}Cong)$, $(TApp\text{-}Cong)$.

Basis is immediate: the cases of (β^1), (β^2) are guaranteed by Lemmas 14.1 and 14.3.

Inductive Hypothesis. Suppose the claim of the Theorem is true for all reduction judgments appearing in the proof of the reduction judgment $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$, except itself. We have to prove that the claim holds for $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ itself. It suffices to make the case analysis on a last rule from Definition 15.1 applied in the proof of $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$.

Case of $(Trans)$**:** Let $\Gamma \vdash_n^{(k)} t : \theta$ (A) and the reduction judgment $\Gamma \vdash^{(k)} t \rightsquigarrow t' : \theta$ be obtained from $\Gamma \vdash^{(k)} t \rightsquigarrow t'' : \theta$ (B) and $\Gamma \vdash^{(k)} t'' \rightsquigarrow t' : \theta$ (C) by $(Trans)$. Then, from (A) and (B), by inductive hypothesis, $\Gamma \vdash_n^{(k)} t'' : \theta$ (D). Therefore, from (D) and (C), again by inductive hypothesis, $\Gamma \vdash_n^{(k)} t' : \theta$, which is needed.

Case of $(\lambda\text{-}Cong)$**:** Let $\Gamma \vdash_n^{(k)} t \equiv (\lambda x : \sigma\,.\,r) : \sigma \to \tau \equiv \theta$ and the reduction judgment $\Gamma \vdash^{(k)} (\lambda x : \sigma\,.\,r) \rightsquigarrow (\lambda x : \sigma\,.\,r') \equiv t' : \theta$ be obtained from $\Gamma, x : \sigma \vdash^{(k)} r \rightsquigarrow r' : \tau$ (A) by $(\lambda\text{-}Cong)$. From (A), by Auxiliary Lemma 15.2, $\Gamma, x : \sigma \vdash_n^{(k)} r : \tau$ (B). From (B) and (A), by inductive hypothesis, $\Gamma, x : \sigma \vdash_n^{(k)} r' : \tau$ (C). From (C), by Normalization Lemma 5.1, $\Gamma \vdash_n^{(k)} t' \equiv (\lambda x : \sigma\,.\,r') : \sigma \to \tau \equiv \theta$, which is needed.

Case of $(FApp\text{-}Cong\text{-}L)$**:** Let $\Gamma \vdash_n^{(k)} t \equiv r\,s : \tau$ and $\Gamma \vdash^{(k)} r\,s \rightsquigarrow r'\,s \equiv t' : \tau$ (A) be derived from $\Gamma \vdash^{(k)} r \rightsquigarrow r' : \sigma \to \tau$ (B) and $\Gamma \vdash_n^{(k)} s : \sigma$ (C) by $(FApp\text{-}Cong\text{-}L)$. From (B), by Auxiliary Lemma 15.2, $\Gamma \vdash_n^{(k)} r : \sigma \to \tau$ (D). By inductive hypothesis, (D) and (B) imply $\Gamma \vdash_n^{(k)} r' : \sigma \to \tau$ (E). From (E) and (C), by Normalization Lemma 5.1, $\Gamma \vdash_n^{(k)} t' \equiv r'\,s : \tau$, which is needed.

Case of $(FApp\text{-}Cong\text{-}R)$**:** Let $\Gamma \vdash_n^{(k)} t \equiv r\,s : \tau$ and $\Gamma \vdash^{(k)} r\,s \rightsquigarrow r\,s' \equiv t' : \tau$ (A) be derived from $\Gamma \vdash_n^{(k)} r : \sigma \to \tau$ (B) and $\Gamma \vdash^{(k)} s \rightsquigarrow s' : \sigma$ (C) by $(FApp\text{-}Cong\text{-}R)$. From (C), by Auxiliary Lemma 15.2, $\Gamma \vdash_n^{(k)} s : \sigma$ (D). By inductive hypothesis, (D) and (C) imply $\Gamma \vdash_n^{(k)} s' : \sigma$ (E). From (B) and (E), by Normalization Lemma 5.1, $\Gamma \vdash_n^{(k)} t' \equiv r\,s' : \tau$, which is needed.

Case of $(\Lambda\text{-}Cong)$**:** Let $\Gamma \vdash_n^{(k)} t \equiv (\Lambda\alpha \le \sigma\,.\,r) : (\forall\alpha \le \sigma\,.\,\tau) \equiv \theta$ and the judgment $\Gamma \vdash^{(k)} (\Lambda\alpha \le \sigma\,.\,r) \rightsquigarrow (\Lambda\alpha \le \sigma\,.\,r') \equiv t' : \theta$ be obtained from $\Gamma, \alpha \le \sigma \vdash^{(k)} r \rightsquigarrow r' : \tau$ (A) by $(\Lambda\text{-}Cong)$. From (A), by Auxiliary Lemma 15.2, $\Gamma, \alpha \le \sigma \vdash_n^{(k)} r : \tau$ (B). From (B) and (A), by inductive hypothesis, $\Gamma, \alpha \le \sigma \vdash_n^{(k)} r' : \tau$ (C). From (C), by Normalization Lemma 5.1, $\Gamma \vdash_n^{(k)} t' \equiv (\Lambda\alpha \le \sigma\,.\,r') : (\forall\alpha \le \sigma\,.\,\tau) \equiv \theta$, which is needed.

Case of (*TApp-Cong*)**:** Let $\Gamma \vdash_n^{(k)} t \equiv r\{\rho\} : \tau[\rho/\alpha] \equiv \theta$ and the refuction judgment $\Gamma \vdash^{(k)} r\{\rho\} \rightsquigarrow r'\{\rho\} \equiv t' : \tau[\rho/\alpha]$ be derived from $\Gamma \vdash^{(k)} r \rightsquigarrow r' : (\forall\alpha \leq \sigma\,.\,\tau)$ (A) and $\Gamma \vdash^{(k)} \rho \leq \sigma$ (B) by (*TApp-Cong*). From (A), by Auxiliary Lemma 15.2,
$\Gamma \vdash_n^{(k)} r : (\forall\alpha \leq \sigma\,.\,\tau)$ (C). By inductive hypothesis, (C) and (A) imply
$\Gamma \vdash_n^{(k)} r' : (\forall\alpha \leq \sigma\,.\,\tau)$ (D). From (D) and (B), by Normalization Lemma 7.1 (and because of (29)) we get $\Gamma \vdash_n^{(k)} t' \equiv r'\,s : \tau$, which is needed. □

17 B. Pierce's Counterexample

Benjamin Pierce suggested the following elegant "counterexample", well illustrating the above notions and problems. It also shows that the safety preconditions in Subject Reduction Lemmas are essential and could not be relaxed.

Consider the following types and judgments defined in Section 15 of [Vor94b]:

$$\begin{aligned}
\tau &\equiv \forall\gamma \leq \top\,.\,((\top \to \top) \to \top) \\
\rho &\equiv \forall\gamma \leq \top\,.\,(\top \to \top) \\
&\quad \Gamma,\, \alpha \leq \tau \vdash^{(0)} \alpha \leq \rho \qquad (30) \\
&\quad \Gamma \vdash^{(0)} \tau \leq \rho \qquad (31)
\end{aligned}$$

Proposition 17.1 *For every* **SnS***-based [Vor95] set of* $\forall$*-hypotheses* $\mathcal{H}$ *the judgment* (30) *is true and the judgment* (31) *is false.* □

This is established by direct calculations. (Similarly, as shown in Section 16 of [Vor94b], one can find judgments with the same properties at every level $\vdash^{(k)}$, $k \in \mathbb{N}$).

Let Γ be the context $i : \top$. Define the following terms:

$$\begin{array}{llll}
u & \equiv_{df} & \lambda g : \rho\,.\,g\{\top\}\,i & : \quad \rho \to \top \\
v & \equiv_{df} & \Lambda\alpha \leq \tau\ \lambda h : \alpha\,.\,u\,h & : \quad \forall\alpha \leq \tau\,.\,(\alpha \to \top) \\
f & \equiv_{df} & \Lambda\gamma \leq \top\ \lambda g : \top \to \top\,.\,g\,i & : \quad \forall\gamma \leq \top\,.\,((\top \to \top) \to \top) \equiv \tau \\
w & \equiv_{df} & (v\{\tau\})\,f & : \quad \top
\end{array}$$

(the right column contains their minimal types in context Γ computed immediately by the rules of Section 4, see also the typing proofs below).

Consider the β-reduction chain for the term w:

$$(v\{\tau\})\,f \;\to\; (\lambda h : \tau\,.\,u\,h)\,f \;\to\; (\lambda g : \rho\,.\,g\{\top\}\,i)\,f \;\to\; f\{\top\}\,i \;\to\; i\,i \qquad (32)$$

The last term $i\,i$ in this sequence is *ill-typed*. What is the problem?

Let us consider the typing proof of the first (β^2)-redex in (32):

$$v\{\tau\} \;\equiv\; (\Lambda\alpha \leq \tau\,.\,\lambda h : \alpha\,.\,((\lambda g : \rho\,.\,g\,\top\,i)\,h))\{\tau\} \qquad (33)$$

For typographical reasons we have to cut the proof into pieces.

Subproof 1:

$$\dfrac{\dfrac{\dfrac{\Gamma,\, \alpha \leq \tau,\, h : \alpha,\, g : \rho \;\vdash^{(0)}_{mn}\; g : \rho \quad \ldots \vdash^{(0)}\; \top \leq \top}{\Gamma,\, \alpha \leq \tau,\, h : \alpha,\, g : \rho \;\vdash^{(0)}_{mn}\; g\,\top : \top \to \top}\,(\mu\text{-}TApp) \qquad \ldots \vdash^{(0)}_{n}\; i : \top}{\Gamma,\, \alpha \leq \tau,\, h : \alpha,\, g : \rho \;\vdash^{(0)}_{mn}\; g\,\top\, i : \top}\,(\mu\text{-}FApp)}{\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{mn}\; (\lambda g : \rho\,.\, g\,\top\, i) : \rho \to \top}\,(\mu\text{-}FAbs)$$

where $\ldots \vdash^{(0)}_{n}\; i : \top$ is proved from the axioms $\ldots \vdash^{(0)}_{mn}\; i : \top$ (recall that $\Gamma \;\equiv\; i : \top$) and $\ldots \vdash^{(0)}\; \top \leq \top$ by (Sub).

Subproof 2:

$$\dfrac{\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{mn}\; h : \alpha \qquad \Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}\; \alpha \leq \rho}{\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{n}\; h : \rho}\,(Sub)$$

The right subtyping premise in this proof is equivalent to (30).

Subproof 3:

$$\dfrac{\dfrac{\text{(*See Subproof 1*)} \qquad \text{(*See Subproof 2*)}}{\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{mn}\; (\lambda g : \rho\,.\, g\,\top\, i) : \rho \to \top \qquad \Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{n}\; h : \rho}\,(\mu\text{-}FApp)}{\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}_{mn}\; (\lambda g : \rho\,.\, g\,\top\, i)\, h : \top}\,(\mu\text{-}FAbs)$$

$$\Gamma,\, \alpha \leq \tau \;\vdash^{(0)}_{mn}\; (\lambda h : \alpha\,.\, ((\lambda g : \rho\,.\, g\,\top\, i)\, h)) : \alpha \to \top$$

The typing proof for the (β^2)-redex (33) finishes by:

$$\dfrac{\dfrac{\text{(*See Subproof 3*)}}{\Gamma,\, \alpha \leq \tau \;\vdash^{(0)}_{mn}\; (\lambda h : \alpha\,.\, ((\lambda g : \rho\,.\, g\,\top\, i)\, h)) : \alpha \to \top}\,(\mu\text{-}TAbs) \qquad \Gamma \;\vdash^{(0)}\; \tau \leq \tau}{\Gamma \;\vdash^{(0)}_{mn}\; (\Lambda\alpha \leq \tau\,.\, \lambda h : \alpha\,.\, ((\lambda g : \rho\,.\, g\,\top\, i)\, h))\; \{\tau\} : \tau \to \top}\,(\mu\text{-}TApp)$$

The **explanation** to the above "paradox" is therefore simple: the topmost judgment

$$\Gamma \;\vdash^{(0)}_{mn}\; \Lambda\alpha \leq \tau\,.\, \lambda h : \alpha\,.\, ((\lambda g : \rho\,.\, g\,\top\, i)\, h) : (\forall\alpha \leq \tau\,.\, (\alpha \to \top))$$

in this last proof is **not** $[\tau/\alpha]$**-safe**, because the subtyping judgment

$$\Gamma,\, \alpha \leq \tau,\, h : \alpha \;\vdash^{(0)}\; \alpha \leq \rho,$$

occurring in its proof (see the right premise in Subproof 2) is **not** $[\tau/\alpha]$**-stable**. In fact, substituting $[\tau/\alpha]$ into this judgment yields:

$$\Gamma,\, h : \tau \;\vdash^{(0)}\; \tau \leq \rho$$

and this last subtyping judgment, equivalent to (31), is false by the choice of types τ and ρ.

Thus, the failure of the well-typedness preservation for the chain (32) is explained by the violation of the safety precondition of Lemma 14.3.

References

[BL90] K. B. Bruce and G. Longo. A modest model of records, inheritance and bounded quantification. *Information and Computation*, 87:196–240, 1990.

[BTCCS91] V. Breazu-Tannen, T. Coquand, Gunter C., and A. Scedrov. Inheritance as implicit coercion. *Mathematical Structures in Computer Science*, 93:172–221, 1991.

[CG92] P.-L. Curien and G. Ghelli. Coherence of subsumption, minimum typing, and type checking in $F_{\leq}$. *Mathematical Structures in Computer Science*, 2:55–91, 1992.

[CMMS94] L. Cardelli, S. Martini, J.C. Mitchell, and A. Scedrov. An extension of system F with subtyping. *Information and Computation*, 1994. To appear.

[CP94] G. Castagna and B. C. Pierce. Decidable bounded quantification. In *21st ACM Symp. on Principles of Programming Languages (POPL'94)*, pages 151–162, 1994.

[CW85] L. Cardelli and P. Wegner. On understanding types, data abstraction, and polymorphism. *Computing Surveys*, 17(4):471–522, 1985.

[Pie92] B. C. Pierce. Bounded quantification is undecidable. In *19th ACM Symp. on Principles of Programming Languages (POPL'92)*, pages 305–315, 1992.

[Rab69] M. Rabin. Decidability of second order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 141:1–35, 1969.

[TMR53] A. Tarski, A. Mostowski, and R. M. Robinson. *Undecidable theories*. Studies in Logic and the Foundations of Mathematics. North-Holland Publishing Company, 1953. Third printing, 1971.

[Vor94a] S. Vorobyov. $F_{\leq}$: Bounded quantification is NOT essentially undecidable. Technical Report CRIN–94–R–018, Centre de Recherche en Informatique de Nancy, January 1994.

[Vor94b] S. Vorobyov. Hierarchies of decidable extensions of bounded quantification. Research Report INRIA–RR–2354, Centre de Recherche en Informatique de Nancy, September 1994.

[Vor94c] S. Vorobyov. Extensions of $F_{\leq}$ with Decidable Typing. Technical Report CRIN–94–R–127, Centre de Recherche en Informatique de Nancy, September 1994.

[Vor95] S. Vorobyov. Structural decidable extensions of bounded quantification. In *22nd ACM Symp. on Principles of Programming Languages (POPL'95)*, 1995. Also: Technical Report MPI–I–94–257, Max-Planck-Institut für Informatik, Saarbrücken.

Note. All papers of the author are available through `http://www.mpi-sb.mpg.de/guide/staff/sv/sv.html`

Below you find a list of the most recent technical reports of the research group *Logic of Programming* at the Max-Planck-Institut für Informatik. They are available by anonymous ftp from our ftp server `ftp.mpi-sb.mpg.de` under the directory `pub/papers/reports`. Most of the reports are also accessible via WWW using the URL `http://www.mpi-sb.mpg.de`. If you have any questions concerning ftp or WWW access, please contact `reports@mpi-sb.mpg.de`. Paper copies (which are not necessarily free of charge) can be ordered either by regular mail or by e-mail at the address below.

Max-Planck-Institut für Informatik
Library
attn. Regina Kraemer
Im Stadtwald
D-66123 Saarbrücken
GERMANY
e-mail: `kraemer@mpi-sb.mpg.de`

MPI-I-95-2-005	F. Baader, H.-J. Ohlbach	A Multi–Dimensional Terminological Knowledge Representation Language
MPI-I-95-2-002	H. J. Ohlbach, R. A. Schmidt	Functional Translation and Second–Order Frame Properties
MPI-I-95-2-001	S. Vorobyov	Proof normalization and subject reduction in extensions of Fsub
MPI-I-94-261	P. Barth, A. Bockmayr	Finite Domain and Cutting Plane Techniques in CLP($\mathcal{PB}$)
MPI-I-94-257	S. Vorobyov	Structural Decidable Extensions of Bounded Quantification
MPI-I-94-254		Report and abstract not published
MPI-I-94-252	P. Madden	A Survey of Program Transformation With Special Reference to *Unfold/Fold* Style Program Development
MPI-I-94-251	P. Graf	Substitution Tree Indexing
MPI-I-94-246	M. Hanus	On Extra Variables in (Equational) Logic Programming
MPI-I-94-241	J. Hopf	Genetic Algorithms within the Framework of Evolutionary Computation: Proceedings of the KI-94 Workshop
MPI-I-94-240	P. Madden	Recursive Program Optimization Through Inductive Synthesis Proof Transformation
MPI-I-94-239	P. Madden, I. Green	A General Technique for Automatically Optimizing Programs Through the Use of Proof Plans
MPI-I-94-238	P. Madden	Formal Methods for Automated Program Improvement
MPI-I-94-235	D. A. Plaisted	Ordered Semantic Hyper-Linking
MPI-I-94-234	S. Matthews, A. K. Simpson	Reflection using the derivability conditions
MPI-I-94-233	D. A. Plaisted	The Search Efficiency of Theorem Proving Strategies: An Analytical Comparison
MPI-I-94-232	D. A. Plaisted	An Abstract Program Generation Logic
MPI-I-94-230	H. J. Ohlbach	Temporal Logic: Proceedings of the ICTL Workshop
MPI-I-94-229	Y. Dimopoulos	Classical Methods in Nonmonotonic Reasoning