ARE THERE HILBERT-STYLE PURE TYPE SYSTEMS?

MARTIN W. BUNDER [a] AND WIL J. M. DEKKERS [b]

[a] School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia
e-mail address: martin_bunder@uow.edu.au

[b] Department of Computer Science, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands
e-mail address: wil@cs.ru.nl

ABSTRACT. For many a natural deduction style logic there is a Hilbert-style logic that is equivalent to it in that it has the same theorems (i.e. valid judgements $\Gamma \vdash P$ where $\Gamma = \emptyset$). For intuitionistic implicational logic, the axioms of the equivalent Hilbert-style logic can be propositions which are also known as the types of the combinators **I**, **K** and **S**.

Natural deduction versions of illative combinatory logics have formulations with axioms that are actual type statements for **I**, **K** and **S**. As pure type systems (PTSs) are, in a sense, equivalent to systems of illative combinatory logic, it might be thought that Hilbert style PTSs (HPTSs) could be based in a similar way.

This paper shows that some PTSs have very trivial equivalent HPTSs, with only the axioms as theorems and that for many PTSs no equivalent HPTSs can exist. Most commonly used PTSs belong to these two classes.

For some PTSs however, including λ^* and the PTS at the basis of the proof assistant Coq, there is a nontrivial equivalent HPTS, with axioms that are type statements for **I**, **K** and **S**.

INTRODUCTION

Most early logical systems (for propositional and predicate logic) allowed no hypotheses and so had no rules for introducing or cancelling them. These could be represented by a finite set of axiom schemes and rules of inference such as modus ponens and generalisation.

Later natural deduction systems which did allow hypotheses had fewer axiom schemes but required introduction and elimination rules for hypotheses. Herbrand showed that classical Hilbert style and natural deduction style propositional and predicate logics had the same theorems (i.e. judgements with empty contexts).

Pure type systems (PTSs), defined below, have two rules that introduce hypotheses and two that cancel them. In this paper we answer a question of Fairouz Kamareddine "Are there Hilbert style PTSs?". When we define Hilbert style PTSs (HPTSs) as PTSs with empty contexts, with a finite set of extra axiom schemes, with $s_1, s_2, s_3, \ldots$ representing arbitrary sorts, and some extra rules, it is obvious that there are HPTSs. We will be interested in whether, for PTSs, there are theorem equivalent HPTSs. We will answer this question for a number of classes of PTSs which include all the PTSs, from the standard literature, that we have examined. The methods we use, for proving that a HPTS is equivalent to a PTS, are along the lines of those of Herbrand, but rather more complex.

Just as combinator based programming languages, requiring no free, or in fact, no variables, have proved useful in practice, perhaps an HPTS, which also requires no (free) variables, that is theorem equivalent to a PTS may be useful. Also, perhaps some metatheoretical results may be proved more easily for an HPTS than for the equivalent PTS.

1. Pure Type Systems

Each Pure Type System (PTS) λX has a set of variables V, a set of constants $\mathcal{C}$, a set of "sorts" $\mathcal{S} \subseteq \mathcal{C}$. It has a class of pseudoterms given by $\mathcal{T} = V \,|\, \mathcal{C} \,|\, (\Pi V : \mathcal{T}.\mathcal{T}) \,|\, (\lambda V : \mathcal{T}.\mathcal{T}) \,|\, \mathcal{T}\mathcal{T}$. If M and N are pseudoterms, $M : A$ is a statement, Γ is a context if it is a sequence of statements; $\Gamma \vdash M : A$ is then called a judgement. A PTS has a set of axioms $\mathcal{A}$ each of the form $c : s$ where $c \in \mathcal{C}$ and $s \in \mathcal{S}$. Then it has a set $\mathcal{R}$ of triples $(s_1, s_2, s_3) \in \mathcal{S}^3$, which determine under what conditions a term $\Pi x : A.B$ is in a sort. Most PTSs are known by a "specification" $(\mathcal{S}, \mathcal{A}, \mathcal{R})$ (as usually $\mathcal{C} = \mathcal{S}$).

The PTS postulates are as follows:

$$\text{(axiom)} \qquad \frac{c : s \in \mathcal{A}}{\vdash c : s}$$

$$\text{(start)} \qquad \frac{\Gamma \vdash A : s}{\Gamma, x : A \vdash x : A}$$

$$\text{(weakening)} \qquad \frac{\Gamma \vdash M : B \qquad \Gamma \vdash A : s \qquad x \notin FV(\Gamma)}{\Gamma, x : A \vdash M : B}$$

$$\text{(application)} \qquad \frac{\Gamma \vdash M : (\Pi x : A.B) \qquad \Gamma \vdash N : A}{\Gamma \vdash MN : B[x := N]}$$

$$\text{(abstraction)} \qquad \frac{\Gamma, x : A \vdash M : B \qquad \Gamma \vdash (\Pi x : A.B) : s}{\Gamma \vdash (\lambda x : A.M) : (\Pi x : A.B)}$$

$$\text{(product)} \qquad \frac{\Gamma, x : A \vdash B : s_2 \qquad \Gamma \vdash A : s_1 \qquad (s_1, s_2, s_3) \in \mathcal{R}}{\Gamma \vdash (\Pi x : A.B) : s_3}$$

$$\text{(conversion)} \qquad \frac{\Gamma \vdash M : A \qquad \Gamma \vdash B : s \qquad A =_\beta B}{\Gamma \vdash M : B}.$$

When there are two judgements as premises in a rule, we call the left one the *major premise* and the right one the *minor premise*.

Later we will need the following definition:

Definition 1.1 (Inhabited and Normal Form Inhabited Sorts)**.**
  s is an *inhabited sort* $(s \in \mathcal{I})$ if $\vdash A : s$ for some A.
  s is a *normal form inhabited sort* $(s \in \mathcal{N})$ if for some term A in normal form, $\vdash A : s$.

The translation $[\ ]$ of Bunder and Dekkers [3] translates the pseudoterms and statements of PTSs into terms of illative combinatory logic (ICL) as follows:

$$[x] = x, \quad [c] = c, \quad [XY] = [X][Y]$$

$$[X : A] = [A][X], \quad [\Pi x{:}X.Y\,x] = \mathbf{G}[X][Y] \quad (x \notin FV(XY))$$

where $\mathbf{G} = \lambda xyz.\Xi x(\mathbf{S}yz)$ ($\mathbf{S}$ is the combinator equivalent to $\lambda xyz.xz(yz)$). Terms in ICL can be represented without any free variables at all using the combinators $\mathbf{S}$ and $\mathbf{K}$ (equivalent to $\lambda xy.x$). Ξxy represents roughly $(\forall u \in x)y(u)$ or $x \subseteq y$.

ICL, designed as a foundation for logic and mathematics, has a rule like (abstraction) which was derived in Bunder [2] from a set of axioms. In Section 6 we will see how the methods developed there lead to the ones used here. The main difference between PTSs and standard ICLs, other than the lack of distinction between terms and types, lies in the (abstraction) rule. The direct counterpart to the ICL rule would have $\Gamma \vdash A : s$, for $\Gamma \vdash (\Pi x{:}A.B) : s$. This is the most important factor in making it difficult to have equivalent Hilbert-style PTSs.

2. HILBERT-STYLE PTSs

We define Hilbert-style PTSs as follows:

Definition 2.1 (HPTS)**.** Each *Hilbert style Pure Type System* (HPTS) has V, $\mathcal{C}$, $\mathcal{S}$, $\mathcal{T}$, statements, contexts and judgements as for PTSs, except that the contexts are always empty. A HPTS has a set of sorts $\mathcal{S}$ and a set of axioms $\mathcal{A}$, as for PTSs, and an additional finite set $\mathcal{B}$ of axiom schemes in which "sort variables" can be replaced by sorts. Most HPTSs are known by a "specification" $(\mathcal{S}, \mathcal{A}, \mathcal{B})$ (as usually $\mathcal{C} = \mathcal{S}$). A HPTS has the PTS (application) and (conversion) rules (with empty contexts) as well as:

$$\text{(type reduction)} \qquad \frac{\vdash M : A \qquad A \to_\beta B}{\vdash M : B.}$$

$$\text{(subject reduction)} \qquad \frac{\vdash M : A \qquad M \to_\beta N}{\vdash N : A.}$$

Note the latter rules are derivable for all PTSs, for HPTSs neither is, even using (conversion).

Definition 2.2 (Equivalent HPTS)**.** If λX is a PTS with specification $(\mathcal{S}, \mathcal{A}, \mathcal{R})$, a HPTS λX^h, with specification $(\mathcal{S}, \mathcal{A}, \mathcal{B})$ will be *equivalent* if

$$(\forall M, A)(\vdash^X M : A \;\Leftrightarrow\; \vdash^{X^h} M : A).$$

Here $\vdash^X$ stands for provability in λX and $\vdash^{X^h}$ in λX^h. If the PTS is arbitrary or obvious from the context we use $\vdash$ and $\vdash^h$. $\mathcal{B}$ will a function of $\mathcal{R}$, i.e. it will include axioms such as $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2).\Pi x{:}u.vx] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_3]$ if $(s_1, s_2, s_3) \in \mathcal{R}$.

Below are some PTSs that have been studied in the literature (particularly Barendregt [1] and Geuvers [4]).

In $\lambda^\tau, \mathcal{S} = \{*\}$, $\mathcal{C} = \{*, 0\}$, in all other cases $\mathcal{C} = \mathcal{S}$ consists of all the constants visible in $\mathcal{A}$ and $\mathcal{R}$. (s_1, s_2) is used as an abbreviation for (s_1, s_2, s_2).

$\lambda^\rightarrow$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *)\}$

λ^τ $\mathcal{A} = \{0 : *\}$ $\mathcal{R} = \{(*, *)\}$

λ^* $\mathcal{A} = \{* : *\}$ $\mathcal{R} = \{(*, *)\}$

$\lambda 2$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (\Box, *)\}$

λP $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (*, \Box)\}$

$\lambda\underline{\omega}$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (\Box, \Box)\}$

$\lambda\omega$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (\Box, *), (\Box, \Box)\}$

$\lambda P2$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (\Box, *), (*, \Box)\}$

$\lambda P\underline{\omega}$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (*, \Box), (\Box, \Box)\}$

$\lambda P\omega = \lambda C$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (*, \Box), (\Box, *), (\Box, \Box)\}$

$\lambda\text{AUT-68}$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (*, \Box, \triangle), (\Box, *, \triangle), (\Box, \Box, \triangle), (*, \triangle), (\Box, \triangle)\}$

$\lambda\text{AUT-QE}$ $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *), (*, \Box), (\Box, *, \triangle), (\Box, \Box, \triangle), (*, \triangle), (\Box, \triangle)\}$

λPAL $\mathcal{A} = \{* : \Box\}$ $\mathcal{R} = \{(*, *, \triangle), (*, \Box, \triangle), (\Box, *, \triangle), (\Box, \Box, \triangle), (*, \triangle), (\Box, \triangle)\}$

λU $\mathcal{A} = \{* : \Box, \Box : \triangle\}$ $\mathcal{R} = \{(*, *), (\Box, *), (\Box, \Box), (\triangle, \Box), (\triangle, *)\}$

λHOL $\mathcal{A} = \{* : \Box, \Box : \triangle\}$ $\mathcal{R} = \{(*, *), (\Box, *), (\Box, \Box)\}$

The PTS used in the proof assistant Coq we will call λCoq. It has as axioms:

$$\vdash *_p : \Box_1 \quad \vdash *_s : \Box_1 \quad \forall i \in I\!\!N \vdash \Box_i : \Box_{i+1}.$$

More axioms are generated by

$$A : B, B : C \in \mathcal{A} \;\Rightarrow\; A : C \in \mathcal{A}.$$

In early versions $\mathcal{R}$ is given by

$$(*_s, *_s), (*_p, *_p), (*_s, *_p), (*_p, *_s), (*_p, \Box_i), (*_s, \Box_i), (\Box_i, *_p), (\Box_i, *_s), (\Box_i, \Box_j, \Box_{max(i,j)}) \in \mathcal{R}.$$

for all $i, j \in I\!\!N$. Coq 8.0 replaces $(\Box_i, *_s) \in \mathcal{R}$ by $(\Box_i, *_s, \Box_i) \in \mathcal{R}$.

We will be able to determine whether or not there are equivalent HPTSs for all of the above.

3. Some PTS Lemmas and Definitions

We now state a number of standard lemmas for PTSs. Most proofs can be found in Barendregt [1] or Bunder and Dekkers [3].

Lemma 3.1 (Free Variable Lemma). *If $x_1 : A_1, \ldots, x_n : A_n \vdash M : B$, then*

  (i) $x_1, \ldots, x_n$ *are distinct;*
 (ii) $FV(M, B) \subseteq \{x_1, \ldots, x_n\}$;
(iii) $FV(A_i) \subseteq \{x_1 \ldots, x_{i-1}\}$ *for $1 \leq i \leq n$.* $\qquad\square$

Lemma 3.2 (Substitution Lemma). *If $\Gamma_1, x : A, \Gamma_2 \vdash M : B$ and $\Gamma_1 \vdash N : A$ then $\Gamma_1, \Gamma_2[x := N] \vdash M[x := N] : B[x := N]$.* $\qquad\square$

Lemma 3.3 (Condensing Lemma). *If $\Gamma_1, x : A, \Gamma_2 \vdash M : B$, where $x \notin FV(\Gamma_2, M, B)$, then $\Gamma_1, \Gamma_2 \vdash M : B$.* $\qquad\square$

Lemma 3.4 (Generation Lemma). *Let $\Gamma \vdash M : B$. Then*

  (i) $M \equiv c \in \mathcal{C} \;\Rightarrow\; (\exists s \in \mathcal{S}) \; B =_\beta s \;\&\; c : s \in \mathcal{A}$;
 (ii) $M \equiv x \;\Rightarrow\; (\exists C) \; B =_\beta C \;\&\; x : C \in \Gamma$;
(iii) $M \equiv \Pi x{:}C.D \;\Rightarrow\; (\exists s_1, s_2, s_3) \; (s_1, s_2, s_3) \in \mathcal{R} \;\&\; \Gamma \vdash C : s_1$
$$\&\; \Gamma, x : C \vdash D : s_2 \;\&\; B =_\beta s_3;$$
(iv) $M \equiv \lambda x{:}C.N \;\Rightarrow\; (\exists s \in \mathcal{S}) (\exists D) \Gamma, x : C \vdash N : D \;\&\; B =_\beta \Pi x{:}C.D$
$$\&\; \Gamma \vdash (\Pi x{:}C.D) : s;$$
 (v) $M \equiv PQ \;\Rightarrow\; (\exists C, D) \Gamma \vdash P : \Pi x{:}C.D \;\&\; \Gamma \vdash Q : C \;\&\; B =_\beta D[x := Q]$.

In each case the derivations, of the judgements of the form $\Gamma \vdash R : E$ in (iii) *to* (v), *are shorter than that of $\Gamma \vdash M : B$.* $\qquad\square$

Lemma 3.5 (Correctness of Types Lemma). *If $\Gamma \vdash M : B$ then $(\exists s \in \mathcal{S}) \; [B \equiv s$ or $\Gamma \vdash B : s]$.* $\qquad\square$

Lemma 3.6 (Subject and Type Reduction Lemma). *If $\Gamma \vdash M : B$, then*

  (i) $M \twoheadrightarrow_\beta N$ *implies $\Gamma \vdash N : B$,*
 (ii) *and $B \twoheadrightarrow_\beta A$ implies $\Gamma \vdash M : A$.* $\qquad\square$

Lemma 3.7 (Start Lemma). *If $\Gamma \vdash M : B$, then*

  (i) $(c : s) \in \mathcal{A}$ *implies $\Gamma \vdash c : s$,*
 (ii) $\Gamma \equiv x_1 : A_1, \ldots, x_n : A_n$ *implies that for $0 \leq i < n$ there is an $s \in \mathcal{S}$ such that $x_1 : A_1, \ldots, x_i : A_i \vdash A_{i+1} : s$.* $\qquad\square$

4. PTSs where $\mathcal{A}$ is the Set of Theorems

The following lemma specifies a set of PTSs whose axioms are its only theorems. The equivalent HPTS is then trivially one with no extra axioms, i.e. with $\mathcal{B} = \emptyset$.

Lemma 4.1. *In a PTS satisfying*

$$(\forall c, s_1)\big((c : s_1) \in \mathcal{A} \;\Rightarrow\; \sim (\exists s_2, s_3)[(s_1, s_2, s_3) \in \mathcal{R}]\big) \tag{\$}$$

we have $\vdash M : A \;\Leftrightarrow\; M : A \in \mathcal{A}$.

Proof. We show $M : A \in \mathcal{A}$ by induction on the derivation of

$$\vdash M : A. \tag{4.1}$$

(4.1) clearly does not come by (start) or (weakening).

If (4.1) comes by (application) from

$$\vdash P : (\Pi x{:}C.D) \quad \text{and} \quad \vdash Q : C$$

where $M \equiv PQ$ and $A \equiv D[x := Q]$, we have by the induction hypothesis $P : (\Pi x{:}C.D) \in \mathcal{A}$, which is impossible.

If (4.1) comes by (abstraction) from

$$x : B \vdash N : C \quad \text{and} \quad \vdash (\Pi x{:}B.C) : s$$

where $M \equiv \lambda x{:}B.N$ and $A \equiv \Pi x{:}B.C$, then by the induction hypothesis $(\Pi x{:}B.C) : s \in \mathcal{A}$, which is impossible.

If (4.1) comes by (product) from

$$x : B \vdash C : s_2 \quad , \quad \vdash B : s_1 \quad \text{and} \quad (s_1, s_2, s_3) \in \mathcal{R}$$

where $M \equiv \Pi x{:}B.C$ and $A \equiv s_3$, then, by the induction hypothesis, $(B : s_1) \in \mathcal{A}$, which is impossible by (\$).

If (4.1) comes by (conversion) from

$$\vdash M : B \quad , \quad \vdash A : s \quad \text{and} \quad A =_\beta B$$

then by the induction hypothesis $(M : B), (A : s) \in \mathcal{A}$. However then A and B must be in normal form and so $A \equiv B$ and $(M : A) \in \mathcal{A}$.

If (4.1) is an axiom, the result holds trivially. □

This implies the following theorem and corollary.

Theorem 4.2. *A PTS satisfying* (\$) *has an equivalent HPTS, with* $\mathcal{B} = \emptyset$, *but this is trivial in that it has only its axioms as theorems.* □

Corollary 4.3. $\lambda^\rightarrow$ *and* λP *each have an equivalent HPTS, but* $\vdash * : \square$ *is the only theorem of both systems.* □

5. PTSs with no Equivalent HPTS

In $\lambda^\rightarrow$ and λP there is no term A such that $\vdash A : *$ and the only theorem is $\vdash * : \square$.

We can show, by a single (product) rule preceeded by two uses of an axiom and a (start) or (weakening) rule, that in the other PTSs, given in Section 2, there are theorems that are not axioms. Most of these are given below.

Lemma 5.1.

(i) *In* λ^τ *we have* $\vdash (\Pi x{:}0.0) : *$.

(ii) *In* λ^*, $\lambda 2$, $\lambda P2$, $\lambda \omega$, λC, λU *and* λHOL, *we have* $\vdash (\Pi x{:} * .x) : *$.

(iii) *In* $\lambda \underline{\omega}$ *and* $\lambda P\underline{\omega}$, *we have* $\vdash (\Pi x{:} * .*) : \square$.

(iv) *In* λAUT-68, λAUT-QE *and* λPAL *we have* $\vdash (\Pi x{:} * .*) : \triangle$. □

We now give a condition under which, in a PTS, certain sorts have an infinite number of inhabitants of the form $\Pi x{:}A.B$ that are not substitution instances of each other. We show later that many PTSs with this property cannot be equivalent to HPTSs.

Lemma 5.2. *Assume that in a PTS there is a finite sequence $s_1, s_2, \ldots, s_n \in \mathcal{S}$ such that:*

$$(\exists n \in I\!N)\ n > 1\ \&\ s_1 = s_n \in \mathcal{N}\ \&\ (\forall i)\big(1 \le i < n \Rightarrow (\exists s' \in \mathcal{I}\ \&\ (s', s_i, s_{i+1}) \in \mathcal{R})\big) \tag{\$s_1, \ldots, s_n}$$

then

$$\vdash (\Pi x{:}A.B) : s_1$$

for an infinite number of β-distinct terms $\Pi x{:}A.B$ which are not (s_i for s_j) substitution instances of each other.

Proof. Assume that we have ($\$s_1, \ldots, s_n$) for $s_1, s_2, \cdots \in \mathcal{S}$.

As $s_1 \in \mathcal{N}$ we have, for some A_1, in normal form

$$\vdash A_1 : s_1.$$

Now we show, by induction on i that, for $1 < i \le n$, there is a B_{i-1} and an $A_i = \Pi x_{i-1}{:}B_{i-1}.A_{i-1}$ such that

$$\vdash A_i : s_i. \tag{5.1}$$

For each i we have, by ($\$s_1, \ldots, s_n$), an s' such that $(s', s_{i-1}, s_i) \in \mathcal{R}$ and a B_{i-1} such that

$$\vdash B_{i-1} : s'.$$

When $i = 2$ we have $\vdash A_1 : s_1$ above, otherwise we have $\vdash A_{i-1} : s_{i-1}$ by the induction hypothesis. By (weakening) we have

$$x_{i-1} : B_{i-1} \vdash A_{i-1} : s_{i-1}$$

and by (product) we have (5.1). So (5.1) holds for $1 \le i \le n$ and, as we have $s_1 = s_n$,

$$\vdash A_n : s_1.$$

Repeating the above, with A_n for A_1, we get $\vdash A_{2n-1} : s_1$ and similarly $\vdash A_{3n-2} : s_1, \ldots$

If $A_{in-i+1} =_\beta A_{jn-j+1}$ for $i < j$, then

$$\Pi x_{in-i}{:}B_{in-i}.A_{in-i} =_\beta \Pi x_{jn-j}{:}B_{jn-j}.A_{jn-j}$$

and so $A_{in-i} =_\beta A_{jn-j}$ and eventually

$$A_1 =_\beta A_{(j-i)(n-1)+1}.$$

But A_1 is a proper part of $A_{(j-i)(n-1)+1}$ and is in normal form, which is impossible. Hence $A_n, A_{2n-1}, A_{3n-2}, \ldots$ are β-distinct inhabitants of s all of the form $\Pi x{:}A.B$, which are not substitution instances of each other. $\square$

($\$s_1, \ldots, s_n$) is satisfied for many sequences $s_1, s_2, \ldots, s_n$ and many PTSs. Here we list one such sequence and sort for most of the PTSs given in Section 2.

Lemma 5.3.

(i) λ^τ, λ^* *and* $\lambda 2$ *satisfy* ($\$*, *$).
(ii) $\lambda\underline{\omega}$, $\lambda\omega$, $\lambda P2$, $\lambda P\underline{\omega}$, λC, λU *and* λHOL *satisfy* ($\$\square, \square$).
(iii) $\lambda AUT\text{-}68$, $\lambda AUT\text{-}QE$ *and* λPAL *satisfy* ($\$\triangle, \triangle$).

Proof.
  (i) By Lemma 5.1(i), (ii) with $s' = *$.
  (ii) For λP2, by Lemma 5.1(ii), with $s' = *$. For the others with $s' = \square$.
  (iii) By Lemma 5.1(iv) with $s' = \square$. $\square$

Now we can prove the main result in the section.

Theorem 5.4. *If, in a PTS λX, $(\$s_1, \ldots, s_n)$ holds for some $s_1, \ldots, s_n \in \mathcal{S}$ and*

$$(\forall s_1', s_2', s_3')\big((s_1', s_2', s_3') \in \mathcal{R} \Rightarrow (s_1 : s_2') \notin \mathcal{A}\big), \qquad (\$\$s_1)$$

then there is no HPTS equivalent to λX.

Proof. By Lemma 5.2, if $(\$s_1, \ldots, s_n)$ holds we have, for an infinite number of β - distinct terms $\Pi x{:}A.B$, which are not substitution instances of each other

$$\vdash (\Pi x{:}A.B) : s_1.$$

   Suppose that there is an equivalent λX^h.

   As a HPTS has only a finite set of axioms $\mathcal{B}$, at least some must be derived, in λX^h, by (application) and perhaps (conversion), (type reduction) and (subject reduction) from

$$\vdash^h P : (\Pi y{:}D.E) \quad \text{and} \quad \vdash^h Q : D$$

where $PQ \to_\beta \Pi x{:}A.B$ and $E[y := Q] \to_\beta s_1$. By the equivalence of λX and λX^h also:

$$\vdash P : (\Pi y{:}D.E) \quad \text{and} \quad \vdash Q : D. \qquad (5.2)$$

So by correctness of types (Lemma 3.5), for some $s_3' \in \mathcal{S}$

$$\vdash (\Pi y{:}D.E) : s_3'$$

and by the Generation Lemma (Lemma 3.4(iii)) we have:

$$\vdash D : s_1' \quad \text{and} \quad y : D \vdash E : s_2'$$

where $(s_1', s_2', s_3') \in \mathcal{R}$.

   Now by the substitution lemma (Lemma 3.2), (5.2) and, if needed, subject reduction (Lemma 3.6) $\vdash s_1 : s_2'$.

   By Lemma 3.4(i) this contradicts $(\$\$s_1)$, so λX has no HPTS equivalent. $\square$

Theorem 5.5. λ^τ, $\lambda 2$, $\lambda \omega$, $\lambda \underline{\omega}$, λP2, λP$\underline{\omega}$, λC, λAUT-68, λAUT-QE, λPAL, λU *and* λHOL *have no equivalent HPTSs.*

Proof. By Lemma 5.3 and Theorem 5.4. $\square$

   Note that $(\$\$s)$ is not satisfied by any sort s in λ^* and λCoq. For $\lambda^\to$ $(\$\$*)$ holds but $(\$*, s_2, \ldots, *)$ does not for any $s_2, \ldots$ For λP $(\$\$*)$ fails, $(\$\$\square)$ holds but $(\$\square, s_2, \ldots, \square)$ does not for any $s_2, \ldots$

6. How to Prove (abstraction) and (product)

In implicational logic the $\supset$-introduction rule is $\Gamma, A \vdash B \;\Rightarrow\; \Gamma \vdash A \supset B$. The hypothesis A in $\Gamma, A \vdash B$ is cancelled in $\Gamma \vdash A \supset B$. This rule is proved in a Hilbert-style system by induction on the number of steps in a derivation that allows hypotheses. We assume that an hypothesis can be cancelled in the previous step (or steps) and use this to show it can be cancelled in the next. In intuitionistic and classical implicational logic three cases are needed and each requires the Hilbert-style system to have a particular axiom or theorem.

If the hypothesis p is itself the step in the deduction we need

$$\vdash p \supset p.$$

If the deduction step is an axiom or another hypothesis than p we need

$$\vdash q \supset p \supset q.$$

If the deduction step comes by modus ponens from

$$p \vdash q \quad \text{and} \quad p \vdash q \supset r \;,$$

we need

$$\vdash (p \supset q \supset r) \supset (p \supset q) \supset p \supset r.$$

Note that the three theorems we require represent the simple types of the combinators **I**, **K** and **S** (when $\supset$ is replaced by $\rightarrow$).

In illative combinatory logic, the introduction rule for Ξ (restricted generality) is $\Gamma, Ax \vdash Bx \;\Rightarrow\; \Gamma, \mathbf{L}A \vdash \Xi AB$, where $\mathbf{L}$ is a constant, $x \notin FV(\Gamma, A, B)$ and Ax is the hypothesis being cancelled. In the proof of this rule in a Hilbert-style system,(see Bunder[2]), the first two cases are similar to those for the proof of implicational introduction. The third is the case where $\Gamma, Ax \vdash DM$ is derived from $\Gamma, Ax \vdash \Xi CD$ and $\Gamma, Ax \vdash CM$. Again, by induction, we assume that the Ξ-introduction step can be applied to the previous steps.

The axioms of the Hilbert-style system, when rewritten with $U \rightarrow V$ for $FUV \equiv \lambda x.\Xi U(\lambda y.V(xy))$ are:

$$\vdash \ldots [(A \rightarrow A)\,\mathbf{I}]$$
$$\vdash \ldots [(A \rightarrow B \rightarrow A)\,\mathbf{K}]$$
$$\vdash \ldots [((A \rightarrow B \rightarrow C) \rightarrow (A \rightarrow B) \rightarrow A \rightarrow C)\,\mathbf{S}]$$

where $\ldots$ represent conditions involving $\mathbf{L}$ on A, B and C.

These are type assignment statements for **I**, **K** and **S**.

It might be thought that this same technique could be employed for PTSs, using type assignment statements for **I**, **K** and **S**, of the form $\vdash (\ldots \mathbf{I}) : (\ldots A \rightarrow A)$ etc and with hypotheses of the form $x : A$. This however may not work.

If we have a PTS with $(c : s_1) \in \mathcal{A}$ and can prove $x : c \vdash B : s_2$ and/or $x : c \vdash M : B$, perhaps with $M \equiv x$, $c \equiv B$, it may be that (product) and (abstraction) cannot be applied because $(s_1, s_2, s_3) \notin \mathcal{R}$ for any s_3.

This does not mean that $x : c$ can never be cancelled. We may obtain:

$$x : c \vdash P : (\Pi y{:}D.E) \quad \text{and} \quad x : c \vdash Q : D$$

where $x : c$ cannot be cancelled, as, even if we have

$$x : c \vdash (\Pi y{:}D.E) : s_2 \quad \text{and} \quad x : c \vdash D : s_3,$$

(s_1, s_2, s_4) and (s_1, s_3, s_5) may not be in $\mathcal{R}$ for any $s_4, s_5 \in \mathcal{S}$. However if

$$x : c \vdash PQ : E[y := Q] \quad , \quad x : c \vdash E[y := Q] : s_6 \quad \text{and also} \quad (s_1, s_6, s_7) \in \mathcal{R},$$

so that $\vdash (\Pi x{:}c.PQ) : s_7$, we can cancel $x : c$ to give

$$\vdash (\lambda x{:}c.PQ) : (\Pi x{:}c.E[y := Q]).$$

This PTS therefore does have theorems not in $\mathcal{A}$, but it is hard to determine the HPTS corresponding to it.

7. Supersorted PTSs

PTSs that have equivalent HPTSs are λ^* and λCoq (both versions), but these belong to a larger class that has the following property:

Definition 7.1 (Supersorted). A PTS is said to be *supersorted* if:

$$(\forall c \in \mathcal{C})(\exists s \in \mathcal{S}) \ (c : s) \in \mathcal{A} \quad \text{and} \quad (\forall s_1, s_2 \in \mathcal{S}) \ (\exists s_3 \in \mathcal{S}) \ (s_1, s_2, s_3) \in \mathcal{R}.$$

For supersorted PTSs (abstraction) can be simplified.

Theorem 7.2. *In every supersorted PTS (abstraction) can be replaced by:*

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash (\lambda x{:}A.M) : (\Pi x{:}A.B)}.$$

Proof. If

$$\Gamma, x : A \vdash M : B$$

by Lemmas 3.7(ii) and 3.5 we have, for some $s_1, s_2 \in \mathcal{S}$:

$$\Gamma \vdash A : s_1 \quad \text{and either} \quad \Gamma, x : A \vdash B : s_2 \quad \text{or} \quad B = s \text{ for some } s \in \mathcal{S}.$$

If the PTS is supersorted we have, for some s_2, $(s : s_2) \in \mathcal{A}$ in the latter ($B = s$) case, and so the result of the former case by Lemma 3.7(i).

Hence by (product) and supersortedness we have, for some $s_3 \in \mathcal{S}$,

$$\Gamma \vdash (\Pi x{:}A.B) : s_3$$

and by (abstraction) we have

$$\Gamma \vdash (\lambda x{:}A.M) : (\Pi x{:}A.B).$$

For a supersorted PTS λX we define a corresponding HPTS λX^h, which in Theorem 9.4 is shown to be equivalent to λX^h.

Definition 7.3 (Corresponding HPTS). If λX is a supersorted PTS with specification $(\mathcal{S}, \mathcal{A}, \mathcal{R})$ the *corresponding* HPTS λX^h has specification $(\mathcal{S}, \mathcal{A}, \mathcal{B})$, with as members of $\mathcal{B}$ the following theorems of λX:

Axiom Π1 $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2) . \Pi x{:}u.vx] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_3]$ for $(s_1, s_2, s_3) \in \mathcal{R}$.

Axiom I1 $\vdash^h [\lambda x{:}s_1.\lambda y{:}x.y] : [\Pi x{:}s_1.\Pi y{:}x.x]$.

Axiom K1 $\vdash^h [\lambda x{:}s_1.\lambda y{:}s_2.\lambda z{:}x.\lambda u{:}y.z] : [\Pi x{:}s_1.\Pi y{:}s_2.\Pi z{:}x.\Pi u{:}y.x]$.

Axiom S1 $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2).\lambda t{:}(\Pi x{:}u.(\Pi y{:}vx.s_3)).\lambda w{:}(\Pi x{:}u.\Pi y{:}vx.txy).$
$\lambda z{:}(\Pi x{:}u.vx).\lambda x{:}u.wx(zx)] :$
$[\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi t : (\Pi x{:}u.(\Pi y{:}vx.s_3)).\Pi w{:}(\Pi x{:}u.\Pi y{:}vx.txy).$
$\Pi z{:}(\Pi x{:}u.vx).\Pi x{:}u.tx(zx)]$.

and additional axioms of $\mathcal{B}$ generated by (I) and (II):

(I) If $(M : A) \in \mathcal{B}$, $A \notin \mathcal{S}$ and A' is obtained from A by replacing any second occurrence of an s_i in A by any s_j not in A, then if $\vdash A' : s$ for $s \in \mathcal{S}$, $(A' : s) \in \mathcal{B}$. Any conditions on $(M : A) \in \mathcal{B}$ not required in the proof of $\vdash A' : s$ are not part of the new axiom.

(II) If $([\lambda x_1{:}A_1 \ldots \lambda x_{i-1}{:}A_{i-1} \ . \ \Pi x_i{:}A_i.B] : [\Pi x_1{:}A_1 \ldots \Pi x_{i-1}{:}A_{i-1} \ . \ s]) \in \mathcal{B}$ and $s' \in \mathcal{B}$ satisfies $\vdash (\lambda x_1{:}A_1 \ldots \lambda x_i{:}A_i.B) : (\Pi x_1{:}A_1 \ldots \Pi x_i{:}A_i.s')$, then
$$([\lambda x_1{:}A_1 \ldots \lambda x_i{:}A_i.B] : [\Pi x_1{:}A_1 \ldots \Pi x_i{:}A_i.s']) \in \mathcal{B}.$$

Note. The $s_1, s_2, \ldots$ in Axioms I1, K1 and S1 are sort variables that can be replaced by arbitrary elements of $\mathcal{S}$. In the axioms generated by (I) and (II) there are restrictions on the sorts that can be substituted for such variables based on the PTS provability of the judgements mentioned.

Given a PTS λX, we will assume below that λX^h is the corresponding HPTS.

Theorem 7.4. *If, for a supersorted PTS, $\vdash^h M : A$ then $\vdash M : A$.*

Proof. By induction on the derivation of $\vdash^h M : A$.

If $\vdash^h M : A$ is one of the axioms of $\mathcal{A}$, $\Pi1$, I1, K1 or S1, or is generated by (I) or (II), we have $\vdash M : A$.

The (application) and (conversion) cases follow from the induction hypothesis.

The (subject reduction) and (type reduction) cases follow from the induction hypothesis and Lemma 3.6. $\qquad\square$

Lemma 7.5. *In a HPTS corresponding to a supersorted PTS,*

(i) *If $(M : A) \in \mathcal{A} \cup \mathcal{B}$ then there is an $s \in \mathcal{S}$ such that $(A : s) \in \mathcal{A} \cup \mathcal{B}$.*

(ii) *If $([\lambda x_1{:}A_1 \ldots \lambda x_{i-1}{:}A_{i-1} \ . \ \Pi x_i{:}A_i.B] : [\Pi x_1{:}A_1 \ldots \Pi x_{i-1}{:}A_{i-1} \ . \ s]) \in \mathcal{B}$, there is an $s' \in \mathcal{S}$ such that $([\lambda x_1{:}A_1 \ldots \lambda x_i{:}A_i.B] : [\Pi x_1{:}A_1 \ldots \Pi x_i{:}A_i.s']) \in \mathcal{B}$.*

Proof.

(i) If $A \in \mathcal{S}$ this follows by supersortedness.

If $(M : A) \in \mathcal{B}$ and $A \notin \mathcal{S}$, we have $\vdash M : A$ by Theorem 7.4 and $\vdash A : s$, for some $s \in \mathcal{S}$, by Lemma 3.5.

Hence $(A : s) \in \mathcal{B}$ by (I).

(ii) If $([\lambda x_1{:}A_1 \ldots \lambda x_{i-1}{:}A_{i-1} \ . \ \Pi x_i{:}A_i.B] : [\Pi x_1{:}A_1 \ldots \Pi x_{i-1}{:}A_{i-1} \ . \ s]) \in \mathcal{B}$, by Theorem 7.4 and Lemma 3.4(iv) and (iii) we have, for some $s' \in \mathcal{S}$,
$$x_1 : A_1 \ldots x_i : A_i \vdash B : s'.$$

(abstraction) and (II) then give the result. $\qquad\square$

We also need an extension of λX^h that allows hypotheses.

Definition 7.6 (λX^{h+})**.** If λX is a PTS, λX^{h+} has all the postulates of λX^h, also with nonempty contexts, and the (start) and (weakening) rules of λX.

Lemma 7.7. $\vdash^{h+} M : A \Leftrightarrow \vdash^h M : A$.

Proof. Immediate because in a derivation of $\vdash^{h+} M : A$ no (start) or (weakening) rule can be used. No nonempty context can be emptied in λX^{h+}. $\qquad\square$

The extra axioms of $\mathcal{B}$ generated by (I) we will need in the proof of the Correctness of Types Lemma for λX^{h+} (if $\Gamma \vdash^{h+} M : A$ then $\Gamma \vdash^{h+} A : s$ for some s).

Those generated by (II) we need in the proof of (abstraction) to show that, if we have $\Gamma \vdash^{h+} (\Pi x{:}C.D) : s_3$, we also have for $(s_1, s_2, s_3) \in \mathcal{R}$, $\Gamma \vdash^{h+} C : s_1$ and $\Gamma \vdash^{h+} (\lambda x{:}C.D) : (\Pi x{:}C.s_2)$, where the derivation of the latter is no longer than the derivation of $\Gamma \vdash^{h+} (\Pi x{:}C.D) : s_3$. The "no longer than" is needed for proof by induction to work.

Many of the axioms are, in a sense, superfluous. We can for example, prove axioms $\Pi 4, \Pi 8$ and $\Pi 11$ (below) from Axiom K1 and Axiom $\Pi 5$ from Axioms $\Pi 1$ and $\Pi 4$. However, using fewer axioms can mean that the derivation of a $\Gamma \vdash^{h+} (\lambda x{:}C.D) : (\Pi x{:}C.s_2)$ is longer than that of $\Gamma \vdash^{h+} (\Pi x{:}C.D) : s_3$.

To illustrate that the axioms, generated by (I) and (II) above, form finite sets, we list all the ones generated by Axiom $\Pi 1(s_i'$ below is such that $(s_i : s_i') \in \mathcal{A})$. There are another six I axioms, another sixteen K axioms and twentynine more S axioms.

$\Pi 2$ $\vdash^h [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_3] : s_4$.

$\Pi 3$ $\vdash^h [\lambda u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_3] : [\Pi u{:}s_1.s_4]$.

$\Pi 4$ $\vdash^h [\lambda u{:}s_1.s_2] : [\Pi u{:}s_1.s_2']$.

$\Pi 5$ $\vdash^h [\Pi u{:}s_1.s_2] : s_3$.

$\Pi 6$ $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2).\lambda x{:}u.vx] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi x{:}u.s_2]$.

$\Pi 7$ $\vdash^h [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi x{:}u.s_3] : s_4$.

$\Pi 8$ $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2).s_3] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_3']$.

$\Pi 9$ $\vdash^h [\lambda u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi x{:}u.s_3] : [\Pi u{:}s_1.s_4]$.

$\Pi 10$ $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x{:}u.s_2).\Pi x{:}u.s_3] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).s_4]$.

$\Pi 11$ $\vdash^h [\lambda u{:}s_1.\lambda v{:}(\Pi x u.s_2).\lambda x{:}u.s_3] : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi x{:}u.s_3']$.

The axiom required by (I) for $\Pi 1, \Pi 8$ and $\Pi 10$ is $\Pi 2$, for $\Pi 3, \Pi 4$ and $\Pi 9, \Pi 5$, for $\Pi 11$, $\Pi 7$ and for $\Pi 6$ the instance of $\Pi 7$ where $s_2 = s_3$. The axiom required by (II) for $\Pi 1$ is $\Pi 6$, for $\Pi 2$ is $\Pi 3$, for $\Pi 3$ is $\Pi 8$, for $\Pi 5$ is $\Pi 4$, for $\Pi 7$ is $\Pi 9$, for $\Pi 9$ is $\Pi 10$ and for $\Pi 10$ is $\Pi 11$.

Each axiom is an axiom scheme in the sense that it is an axiom for all $s_1, s_2, \ldots$ for which it is provable in λX. Thus most axioms (not $\Pi 6$) have some restrictions, other than $s_i : s_i' \in \mathcal{A}$, for example $(s_1, s_2', s_5), (s_5, s_3', s_4) \in \mathcal{R}$ in $\Pi 3$. Some of these restrictions will appear in (the proofs of) some of the lemmas for λX^{h+} below.

We will show later that in λX^{h+}, for a suitable λX, (product) and (abstraction) are admissible and that the theorems of λX^{h+} are exactly those of λX^h and λX.

8. The Correctness of Types Lemmas for λX^{h+}

To state and prove some preliminary lemmas we need some definitions.

Definition 8.1 (major premise chain)**.** A *major premise chain (mpc)* in a derivation is a sequence of judgements starting with one formed by a (start) rule or an axiom. The remaining judgements of the chain are obtained by (weakening), (application) or (conversion), with the previous judgement as major premise, or by (subject reduction) or (type reduction). The minor premises in (weakening), (application) and (conversion) rules for which the major premises are in an mpc, will be called the *minor premises attached to the mpc.*

The final judgement of an mpc that is not a proper part of a larger mpc, must be the final judgement in a derivation, a judgement that is the premise for a (start) rule or the minor premise in a (weakening), (application) or (conversion) rule.

Any derivation is therefore made up of linked mpcs.

Definition 8.2. An mpc is said to be *long* if it starts with an axiom of the form
$$\vdash^{h+} [\lambda x_1{:}A_1 \ldots \lambda x_n{:}A_n.N] : [\Pi x_1{:}A_1 \ldots \Pi x_n{:}A_n.A_{n+1}] \tag{8.1}$$
where $n > 0$, N is one of $x_1, \ldots, x_n$ or is formed by application from (some of) $x_1, \ldots, x_n$ and the mpc has at least n (application) steps and (subject reduction) steps that reduce all of the n λx_i redexes. An mpc is *short* otherwise.

A derivation is *short* if it has no long mpcs and *long* otherwise.

Definition 8.3 (Application Length - alength)**.** The *application length* or *alength* of a derivation is its number of (application) steps, where steps in identical minor premises in the derivation, are counted only once.

A derivation of lesser alength than another will be called *ashorter*, one of greater alength *alonger*.

Note. One derivation of a judgement may be shorter (in length) than another without being short.

Lemma 8.4. *If the final mpc in a derivation of*
$$\Gamma \vdash^{h+} M : A. \tag{8.2}$$
is long, it starts with an axiom of the form (8.1) *and the (subject reduction) step that reduces the λx_n redex comes directly after the nth (application) step, then that derivation of* (8.2) *can be replaced by an ashorter one.*

Proof. This has to be proved for each of the axioms of $\mathcal{B}$ that is of this form. We will prove it for Axiom S8, below, the proofs for other axioms are similar.

S8 $[\lambda y_1{:}s_q.\lambda y_2{:}(\Pi x{:}y_1.s_r).\lambda y_3{:}(\Pi x{:}y_1.\Pi y{:}y_2 x.s_t).\lambda y_4{:}(\Pi x{:}y_1.\Pi y{:}y_2 x.y_3 xy).$
　　$\lambda y_5{:}(\Pi x{:}y_1.y_2 x).\lambda y_6{:}y_1.y_3 y_6(y_5 y_6)] :$
　　$[\Pi y_1{:}s_q.\Pi y_2{:}(\Pi x{:}y_1.s_r).\Pi y_3{:}(\Pi x{:}y_1.\Pi y{:}y_2 x.s_t).\Pi y_4{:}(\Pi x{:}y_1.\Pi y{:}y_2 x.y_3 xy).$
　　$\Pi y_5{:}(\Pi x{:}y_1.y_2 x).\Pi y_6{:}y_1.s_t)].$

Let the minor premises in the six (application) steps involving Πy_1 to Πy_6 in the long derivation of (8.2) be, for $1 \le i \le 6$:
$$\Gamma_i \vdash^{h+} Y_i : E_i$$
where $E_1 =_\beta s_q$, $E_2 =_\beta \Pi x{:}Y_1.s_r$, $E_3 =_\beta \Pi x{:}Y_1.\Pi y{:}Y_2 x.s_t$, $E_4 =_\beta (\Pi x{:}Y_1.\Pi y{:}Y_2 x.Y_3 xy)$, $E_5 =_\beta \Pi x{:}Y_1.Y_2 x$, $E_6 =_\beta Y_1$, $A =_\beta s_t$ and $Y_3 Y_6(Y_5 Y_6) \twoheadrightarrow_\beta M$.

Then for some Y_1^*, $Y_1 \twoheadrightarrow_\beta Y_1^*$ and $E_6 \twoheadrightarrow_\beta Y_1^*$ and for some $R[x]$, $Y_2 x \twoheadrightarrow_\beta R[x]$, $E_3 \twoheadrightarrow_\beta \Pi x{:}Y_1^*.\Pi y{:}R[x].s_3$ and $E_5 \twoheadrightarrow_\beta \Pi x{:}Y_1^*.R[x]$.

Note that as contexts can only grow, each Γ_i for $1 \le i < 6$ is an initial segment of Γ_{i+1}

Now by (weakening), (subject reduction), (type reduction) and just three (application) steps we get from three of these minor premises:
$$\Gamma_6 \vdash^{h+} Y_3 Y_6 : \Pi y{:}R[X_i].s_t,$$
$$\Gamma_6 \vdash^{h+} Y_5 Y_6 : R[X_i]$$
and so
$$\Gamma_6 \vdash^{h+} Y_3 Y_6(Y_5 Y_6) : s_t$$

which, as $Y_3Y_6(Y_5Y_6) \twoheadrightarrow_\beta M$, gives (8.2).

We now have a new derivation of (8.2), which, given that any (application)s in the two uses of $\Gamma_6 \vdash Y_6 : E_6$ are counted only once, has fewer (application)s, and so is ashorter than, the old derivation of (8.2).

Lemma 8.5 (Shortness Lemma for HPTS$^+$). *Every valid judgement in a HPTS$^+$ has a short derivation.*

Proof. We prove this by showing that for every long derivation there is an ashorter derivation of the same judgement.

Assume that the following is the part of a long mpc, in a long derivation, up to the λx_n reduction, together with the minor premises used in the n (application) steps.

$$\vdash^{h+} T_1 : \Pi x_1{:}B_0.C_0$$

$$\vdots$$

$$\frac{\Gamma_1 \vdash^{h+} T_1 : \Pi x_1{:}B_1.C_1 \qquad\qquad \Gamma_1 \vdash^{h+} X_1 : B_1}{\Gamma_1 \vdash^{h+} T_1X_1 : C_1[x_1 := X_1]}$$

$$\vdots$$

$$\frac{\Gamma_2 \vdash^{h+} T_2 : \Pi x_2{:}B_2.C_2 \qquad\qquad \Gamma_2 \vdash^{h+} X_2 : B_2}{\Gamma_2 \vdash^{h+} T_2X_2 : C_2[x_2 := X_2]}$$

$$\vdots$$

$$\frac{\Gamma_n \vdash^{h+} T_n : \Pi x_n{:}B_n.C_n \qquad\qquad \Gamma_n \vdash^{h+} X_n : B_n}{\Gamma_n \vdash^{h+} T_nX_n : C_n[x_n := X_n]}$$

$$\vdots$$

$$\frac{\Gamma_{n+1} \vdash^{h+} (\lambda x_n{:}A'_n.N')X'_nY_1\dots Y_k : D}{\Gamma_{n+1} \vdash^{h+} N'[x_n := X'_n]Y_1\dots Y_k : D}$$

Here $\vdash^{h+} T_1 : \Pi x_1{:}B_0.C_0$ is an axiom of the form (8.1) with N made up of (some of) $x_1, \dots, x_n$, $T_n \twoheadrightarrow_\beta \lambda x_n{:}A'_n.N'$ and $X_n \twoheadrightarrow_\beta X'_n$. The first, second and nth of the n or more (application)s and the (subject reduction) contracting the λx_n redex are explicitly shown. The steps after the nth (application) only alter T_n by reducing it, so steps can be permuted so that the λx_n reduction takes place straight after the nth (application) step as follows:

$$\frac{\dfrac{\Gamma_n \vdash^{h+} T_n : \Pi x_n{:}B_n.C_n}{\Gamma_n \vdash^{h+} \lambda x_n{:}A'_n.N' : \Pi x_n{:}B_n.C_n} \qquad\qquad \Gamma_n \vdash^{h+} X_n : B_n}{\dfrac{\Gamma_n \vdash^{h+} (\lambda x_n{:}A'_n.N')X_n : C_n[x_n := X_n]}{\dfrac{\Gamma_n \vdash^{h+} N'[x_n := X_n] : C_n[x_n := X_n]}{\vdots}}}$$

$$\Gamma_{n+1} \vdash^{h+} N'[x_n := X'_n]Y_1\dots Y_k : D$$

This new derivation is no alonger than the original, but the part up to $\Gamma_n \vdash^{h+} N'[x_n := X'_n] : C_n[x_n := X_n]$ is long and can be replaced, by Lemma 8.4, by an ashorter derivation, so the whole derivation becomes ashorter. (If the derivation had identical mpcs to the above, which were all minor premises in the same mpc, all would have to be altered as above to ensure that the new derivation is not alonger than the old.)

In the remaining lemmas and theorems we use a different measure of length of a derivation, where "similar" subderivations are counted only once.

Definition 8.6 (Similar)**.** Two derivations are said to be *similar* if they are identical or one, in its final mpc, starts with an axiom of $\mathcal{B}$ of the form (8.1), and the other differs only in that its final mpc starts with an axiom of $\mathcal{B}$ generated from the other by one or more applications of (II).

We now define:

Definition 8.7 (Similarity Length - slength)**.** The *similarity length* (or *slength*) of a derivation is given by:

 (i) the number of (application) steps,
 (ii) the number of (conversion), (start) and (weakening) steps.

Similar derivations ending in the two premises of a (weakening) step, are counted only once.

A derivation of lesser slength than another will be said to be *sshorter* and one of greater slength as *slonger*.

Lemma 8.8. *Given, for $s' \in \mathcal{S}$, a short derivation of:*

$$\Gamma \vdash^{h+} \Pi x{:}B.C : s', \tag{8.3}$$

there is, for some $s \in \mathcal{S}$, a derivation, no longer or slonger than that of (8.3)*, of*

$$\Gamma \vdash^{h+} \lambda x{:}B.C : \Pi x{:}B.s. \tag{8.4}$$

Proof. Consider the first judgement in the final mpc in a short derivation of (8.3). This cannot be an axiom of $\mathcal{A}$ or be formed by a (start) rule, so it is an axiom of $\mathcal{B}$ of the form (8.1), where $N \equiv \Pi x{:}B'.C'$ and $A_{n+1} \equiv s'$.

If in this mpc we replace this axiom by the one generated from it by (II), then using exactly the same steps and minor premises we obtain a derivation of (8.4) of the same length.

In this final mpc there are no (weakening) steps in which the premises are similar, until perhaps after the last (application) step, as, until then, no type can be in $\mathcal{S}$. If, at that stage, (8.3) is formed by one or more (weakening) steps (and perhaps (subject reduction)) from $\Gamma^- \vdash^{h+} \Pi x{:}B^0.C^0 : s'$ and similar minor premises such as $\Gamma^- \vdash^{h+} D : s'$, these are counted only once each in the slength. In the derivation obtained by changing the axiom, the above derivations remain similar and so the slength of the derivation remains the same. $\square$

Lemma 8.9 (Correctness of Types for HPTS$^+$)**.** *If λX is supersorted and*

$$\Gamma \vdash^{h+} M : A, \tag{8.5}$$

then, for some $s \in \mathcal{S}$, $A \equiv s$ or there is a short derivation, of slength no more than that of a short derivation of (8.5)*, of*

$$\Gamma \vdash^{h+} A : s. \tag{8.6}$$

Proof. By induction on the number k, of judgements in the final mpc of a short derivation of $\Gamma \vdash^{h+} M : A$, where $A \notin \mathcal{S}$.

If k=1 and (8.5) comes by a (start) rule from

$$\Gamma^- \vdash^{h+} A : s, \tag{8.7}$$

where $M \equiv x$ and $\Gamma \equiv \Gamma^-, x : A$, (8.6) comes from two copies of (8.7) and (weakening). The two derivations of (8.7) are counted only once, so this derivation of (8.6) is no slonger than that of (8.5).

If (8.5) is an axiom we have (8.6) by (I) or by supersortedness.

We now assume k $\geq$ 2.

If (8.5) comes from $\Gamma' \vdash^{h+} M : A$ and $\Gamma' \vdash^{h+} B : s'$, by (weakening), where $s' \in \mathcal{S}$, $A \notin \mathcal{S}$ and $\Gamma = \Gamma', x : B$, these derivations are both counted in the slength of the derivation of (8.5). We have, by the induction hypothesis, $\Gamma' \vdash^{h+} A : s$, by a derivation no slonger than that of $\Gamma' \vdash^{h+} M : A$, for some $s \in \mathcal{S}$ and we obtain (8.6) by (weakening), by a derivation that is no slonger than that of (8.5).

If (8.5) comes from $\Gamma \vdash^{h+} N : A$, by (subject reduction), we have (8.6) by a derivation no slonger than that of (8.5).

If (8.5) comes from $\Gamma \vdash^{h+} M : B$, by (type reduction), we have $\Gamma \vdash^{h+} B : s$ by the induction hypothesis and (8.6) by (subject reduction), by a derivation no slonger than that of (8.5).

If (8.5) comes from $\Gamma \vdash^{h+} M : B$, by (conversion), we have (8.6) by a derivation sshorter than that of (8.5).

If (8.5) comes from $\Gamma \vdash^{h+} P : \Pi x{:}B.C$ and $\Gamma \vdash^{h+} Q : B$, where $M \equiv PQ$ and $A \equiv C[x := Q]$, by (application), we have by the induction hypothesis, $\Gamma \vdash^{h+} \Pi x{:}B.C : s'$ for some $s' \in \mathcal{S}$, by a derivation no slonger than that of $\Gamma \vdash^{h+} P : \Pi x{:}B.C$. Then by Lemma 8.8 we have $\Gamma \vdash^{h+} \lambda x{:}B.C : \Pi x{:}B.s$, for some $s \in \mathcal{S}$ by a derivation no slonger than that of $\Gamma \vdash^{h+} P : \Pi x{:}B.C$. Then by (application) using $\Gamma \vdash^{h+} Q : B$ we have (8.6) by a derivation no slonger than that of (8.5). $\qquad\square$

Lemma 8.10 (Start Lemma for HPTS$^+$)**.** *If*

$$\Gamma, x : A \vdash^{h+} M : B \tag{8.8}$$

then, for some $s \in \mathcal{S}$,

$$\Gamma \vdash^{h+} A : s.$$

Proof. By an easy induction on the derivation of (8.8). $\qquad\square$

9. The Equivalence Results

Lemma 9.1. *If λX is supersorted, (abstraction) is admissible in λX^{h+}.*

Proof. If λX is supersorted we prove that if

$$\Gamma, x : A \vdash^{h+} M : B \tag{9.1}$$

then

$$\Gamma \vdash^{h+} (\lambda x{:}A.M) : (\Pi x{:}A.B) \tag{9.2}$$

by induction on the slength of a short derivation of (9.1).

Case 1. (9.1) comes by (start) (and (type reduction)) from
$$\Gamma \vdash^{h+} A : s$$
where $M \equiv x$ and $A \twoheadrightarrow_\beta B$.

By Axiom I1 and (application) (and (type reduction)) we have (9.2).

Case 2. (9.1) comes by (weakening) (and reduction) from
$$\Gamma \vdash^{h+} M' : B' \quad \text{and} \quad \Gamma \vdash^{h+} A : s_2$$
then by the Correctness of Types Lemma (Lemma 8.9) or supersortedness, for some s_1.
$$\Gamma \vdash^{h+} B' : s_1$$
and (9.2) follows after three (applications) applied to Axiom K1 (and reduction).

Case 3. (9.1) comes by (conversion) (and reduction) from
$$\Gamma, x : A \vdash^{h+} M : C \quad , \quad \Gamma, x : A \vdash^{h+} B' : s_2 \quad \text{and} \quad C =_\beta B' \twoheadrightarrow_\beta B.$$
By the induction hypothesis and (subject reduction) we have:
$$\Gamma \vdash^{h+} (\lambda x{:}A.M) : (\Pi x{:}A.C) \tag{9.3}$$
and
$$\Gamma \vdash^{h+} (\lambda x{:}A.B) : (\Pi x{:}A.s_2)$$
where $\Pi x{:}A.C =_\beta \Pi x{:}A.B$.

We have by Lemma 8.9 applied to (9.1), for some s_1,
$$\Gamma \vdash^{h+} A : s_1$$
and, by supersortedness $(s_1, s_2, s_3) \in \mathcal{R}$ for some s_3, so by Axiom $\Pi 1$ and (subject reduction),
$$\Gamma \vdash^{h+} (\Pi x{:}A.B) : s_3,$$
hence by (9.3) and (conversion) we have (9.2).

Case 4 (9.1) comes by (application) (and reduction) from
$$\Gamma, x : A \vdash^{h+} P : (\Pi y{:}C.D) \tag{9.4}$$
and
$$\Gamma, x : A \vdash^{h+} Q : C \tag{9.5}$$
where $PQ \twoheadrightarrow_\beta M$ and $D[y := Q] \twoheadrightarrow_\beta B$.

By the Correctness of Types lemma we have for some $s_4 \in \mathcal{S}$, by a derivation no slonger than that of (9.4):
$$\Gamma, x : A \vdash^{h+} (\Pi y{:}C.D) : s_4 \tag{9.6}$$
now by Lemma 8.8 we have for some $s_3 \in \mathcal{S}$, by a derivation no slonger than that of (9.4), and so sshorter than that of (9.4):
$$\Gamma, x : A \vdash^{h+} (\lambda y{:}C.D) : (\Pi y{:}C.s_3) \tag{9.7}$$
Now by the induction hypothesis applied to (9.4), (9.5) and (9.7) we have:
$$\Gamma \vdash^{h+} (\lambda x{:}A.P) : (\Pi x{:}A.\Pi y{:}C.D) \tag{9.8}$$
$$\Gamma \vdash^{h+} (\lambda x{:}A.Q) : (\Pi x{:}A.C) \tag{9.9}$$
$$\Gamma \vdash^{h+} (\lambda x{:}A.\lambda y{:}C.D) : (\Pi x{:}A.\Pi y{:}C.s_3) \tag{9.10}$$

also by the Correctness of Types Lemma applied to (9.9) we have for some $s_4 \in \mathcal{S}$

$$\Gamma \vdash^{h+} (\Pi x{:}A.C) : s_4$$

and by Lemma 8.8 for some $s_2 \in \mathcal{S}$

$$\Gamma \vdash^{h+} (\lambda x{:}A.C) : (\Pi x{:}A.s_2) \tag{9.11}$$

now by Axiom S1, $\Gamma \vdash^{h+} A : s_1$, (obtained as in Case 3) (9.11), (9.10), (9.8), (9.9) and five (application)s, (subject reduction) and (type reduction) give (9.2). (Note that in Axiom S1 s_1, s_2 and s_3 (here s_1, s_5 and s_2) can be arbitrarily chosen in a supersorted PTS). $\qquad\square$

Lemma 9.2. *If λX is supersorted (product) is admissible in λX^{h+}.*

Proof. If $\Gamma \vdash^{h+} A : s_1$, $\Gamma, x : A \vdash^{h+} B : s_2$ and $(s_1, s_2, s_3) \in \mathcal{R}$, by Lemma 9.1,

$$\Gamma \vdash^{h+} \lambda x{:}A.B : \Pi x{:}A.s_2$$

so by Axiom II1, (application) and (subject reduction) we have

$$\Gamma \vdash \Pi x{:}A.B : s_3. \qquad\square$$

Lemmas 9.1 and 9.2 show that a theorem that can be proved in λX^{h+}, using hypotheses, (abstraction) and (product), can also be proved in λX^h. So:

Theorem 9.3. *If λX is supersorted it is equivalent to λX^{h+} in that they have the same valid judgements.*

Proof. By Theorem 7.4, λX^h is a subsystem of λX. The additional rules of λX^{h+} are rules of λX, so λX^{h+} is a subsystem of λX. The extra rules of λX have been shown to be admissible in λX^{h+} in Lemmas 9.1 and 9.2, so λX^{h+} and λX have the same valid judgements. $\qquad\square$

Theorem 9.4. *If λX is supersorted λX and λX^h are equivalent in that they have the same theorems.*

Proof. It follows from Theorem 9.3 that λX and λX^{h+} have the same valid judgements with empty contexts i.e. theorems and so by Lemma 7.7 that λX and λX^h have the same theorems. $\qquad\square$

10. Axioms I, K, S and Π as Types

Axioms I1, K1 and S1 can be rewritten in terms of type free combinators (allowing η-reduction) as:

Axiom I1 $\vdash \mathbf{KI} : (\Pi x{:}s_1.\Pi y{:}x.x)$.

Axiom K1 $\vdash \mathbf{K}(\mathbf{KK}) : (\Pi x{:}s_1.\Pi y{:}s_2.\Pi z : y.\Pi u{:}x.y)$.

Axiom S1 $\vdash \mathbf{K}(\mathbf{K}(\mathbf{KS})) : [\Pi u{:}s_1.\Pi v{:}(\Pi x{:}u.s_2).\Pi t{:}(\Pi x.u.\Pi y{:}vx.s_3).$
$\Pi w{:}(\Pi x{:}u.\Pi y{:}vx.txy).\Pi z{:}(\Pi x{:}u.vx).\Pi x{:}u.tx(zx)]$.

These give the standard types of the combinators (writing $A \to B$ for $\Pi x{:}A.B$ when $x \notin FV(B)$):

$$\vdash A : s_1 \ \Rightarrow\ \vdash \mathbf{I} : A \to A$$

$$\vdash A : s_1, \vdash B : s_2 \ \Rightarrow\ \vdash \mathbf{K} : B \to A \to B$$

$$\vdash A : s_1, \vdash B : A \to s_2, \ \vdash C : (\Pi x{:}A.\Pi y{:}Bx.s_3) \ \Rightarrow\ \vdash \mathbf{S} : \Pi w{:}(\Pi x{:}A.\Pi y{:}Bx.Cxy).$$

$$\Pi z{:}(\Pi x{:}A.Bx).\Pi x{:}A.Cx(zx)$$

or, as a special case

$$\vdash A : s_1, \vdash B' : s_2, \vdash C' : A \to B' \to s_3 \ \Rightarrow \ \vdash \mathbf{S} : (A \to B' \to C') \to (A \to B') \to (A \to C')$$

If $\Pi x{:}U \ . \ V x$, where $x \notin FV(UV)$, were represented as $\mathbf{G}UV$ (as it is in ICL), Axiom $\Pi 1$ represents the type for $\mathbf{G}$ (or $\lambda uv \ . \ \Pi x{:}u \ . \ vx$):

$$\vdash \mathbf{G} : (\Pi u{:}s_1 \ . \ \Pi v{:}(\Pi x{:}u \ . \ s_2) \ . \ s_3) \quad \text{or} \quad \vdash \mathbf{G} : (\Pi u{:}s_1 \ . \ (u \to s_2) \to s_3).$$

11. Identifying λ and Π

In the de Bruijn AUTOMATH systems Π and λ are usually identified. Kameraddine has studied the effect of allowing β-reductions in the (former) Π terms in [5]. Doing this Axiom I1 becomes:

$$\vdash \mathbf{KI} : (\lambda x{:}s_1 \ . \ \lambda y{:}x \ . \ x)$$

and similarly for the other axioms. If we write the type in terms of combinators we can get (depending on the algorithm)

$$\vdash \mathbf{KI} : \mathbf{S(KK)I} \quad \text{or} \quad \vdash \mathbf{KI} : \mathbf{K}.$$

12. Conclusion

We have shown that PTSs come in at least three categories: those satisfying ($\$s$) and ($\$\$s$) that have no equivalent HPTS, those satisfying ($\$s_1 \dots s_n$) that have only a trivial equivalent HPTS and supersorted PTSs, such as λCoq, that have a nontrivial equivalent HPTS. The standard PTSs from the literature that we considered all fit into these categories.

References

[1] Barendregt, H.P. *Lambda calculi with types*, pp. 117-309 in Handbook of Logic in Computer Science, vol 2 of Oxford Science Publications, Oxford University Press, New York, 1992. Theoretical Computer Science, 169:3-21, 1996.

[2] Bunder, M.W. A deduction theorem for restricted generality. *Notre Dame Journal of Formal Logic*, 14:341-346, 1973.

[3] Bunder, M.W. and Dekkers, W.J.M. Equivalences between pure type systems and systems of illative combinatory logics. *Notre Dame Journal of Formal Logic*, 46:181-205, 2005.

[4] Geuvers, H. *Logics and Type Systems*, Thesis University of Nijmegen, 1993.

[5] Kamareddine, F. Typed λ-calculi with one binder. *Journal of Functional Programming*, 15(5):771-796, 2005.