INFORMATION AND COMPUTATION 80, 97–113 (1989)

Unique Normal Forms for Lambda Calculus with Surjective Pairing

J. W. KLOP*

Centre for Mathematics and Computer Science, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands; Department of Mathematics and Computer Science, Free University, de Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

AND

R. C. DE VRIJER

Department of Philosophy, University of Amsterdam, Nieuwe Doelenstraat 15, 1012 CP Amsterdam, The Netherlands; Department of Mathematics and Computer Science, Free University, de Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

We consider the equational theory $\lambda\pi$ of λ-calculus extended with constants π, π_0, π_1 and axioms for surjective pairing: $\pi_0(\pi XY) = X$, $\pi_1(\pi XY) = Y$, $\pi(\pi_0 X)(\pi_1 X) = X$. Two reduction systems yielding the equality of $\lambda\pi$ are introduced; the first is not confluent and, for the second, confluence is an open problem. It is shown, however, that in both systems each term possessing a normal form has a unique normal form. Some additional properties and problems in the syntactical analysis of $\lambda\pi$ and the corresponding reduction systems are discussed.

INTRODUCTION

In this note we consider λ-calculus extended with surjective pairing (SP), that is, extended with constants π, π_0, π_1 and equations $\pi_0(\pi XY) = X$, $\pi_1(\pi XY) = Y$, $\pi(\pi_0 X)(\pi_1 X) = X$. Here π is a pairing operator and π_0, π_1 are projection operators; the third equation amounts to the statement that every object X is a pair—hence the name "surjective" pairing. The equational system λ-calculus plus SP will be denoted here by $\lambda\pi$.

The aim of this paper is threefold. First, we survey several already known results about $\lambda\pi$ and some related systems. Among these results are

* Author partially supported by ESPRIT Project 432: Meteor.

0890-5401/89 $3.00

a counterexample to confluence of a reduction system generating the equality of $\lambda\pi$ (Klop, 1980), and a recent theorem in de Vrijer (1987) stating that $\lambda\pi$ is conservative over λ, the pure λ-calculus.

Second, we present some new results, notably a proof of the fact that two reduction systems which naturally correspond with $\lambda\pi$, have the property of "unique normal forms."

Third, we list some open problems concerning further syntactic properties of the systems under consideration.

The λ-calculus with surjective pairing is of fundamental importance in the theory of categorical logic: every theory with the signature of $\lambda\pi$ and including the axioms of $\lambda\pi$ and the η-axiom $\lambda x \cdot Mx = M$ if x is not free in M, is equivalent in some sense to a certain cartesian closed category called C-monoid. For the precise connection between the category of C-monoids and the category of such extensions of $\lambda\pi$ we refer to Lambek and Scott (1986) (see Corollary 17.6). Recently, work of Curien and others (Curien, 1986, Cousineau, Curien and Mauny, 1985) has shown the relevance of categorical logic for computer science, in particular for implementations of functional languages. (The results of the present paper do not cover the η-axiom, though; see the remarks on open problems in Section 5.)

1. Preliminary Notions

To fix our terminology and notation, we collect in this preliminary section some well-known notions and facts about them. Most of the necessary concepts, such as confluence, can already be defined on an abstract level as follows.

1.1. Definition. (i) An *abstract reduction system* (ARS) is a structure $\mathscr{A} = \langle A, (\rightarrow_\alpha)_{\alpha \in I} \rangle$ consisting of a set A and a sequence of binary relations $\rightarrow_\alpha$ on A, also called (one-step) reduction relations or rewrite relations. Sometimes we will refer to $\rightarrow_\alpha$ as α. If the ARS has only one reduction relation we often drop the subscript. In this paper we will only encounter ARSs having just one reduction relation. (Such structures are called "replacement systems" in Staples, 1975.) If for $a, b \in A$ we have $(a, b) \in \rightarrow_\alpha$, we write $a \rightarrow_\alpha b$ and call b a one-step (α-)reduct of a.

(ii) The transitive reflexive closure of $\rightarrow_\alpha$ is written as $\twoheadrightarrow_\alpha$. So $a \twoheadrightarrow_\alpha b$ if there is a possibly empty, finite sequence of "reduction steps" $a \equiv a_0 \rightarrow_\alpha a_1 \rightarrow_\alpha \cdots \rightarrow_\alpha a_n \equiv b$. Here $\equiv$ denotes identity of elements of A. The element b is called an (α-)reduct of a. The equivalence relation generated by $\rightarrow_\alpha$ is $=_\alpha$, also called the *convertibility* relation generated by $\rightarrow_\alpha$ (or *conversion*).

(iii) The reduction relation $\rightarrow$ is called *weakly confluent* or *weakly Church–Rosser* (WCR) if

$$\forall a, b, c \in A, \exists d \in A \qquad (a \rightarrow b \text{ and } a \rightarrow c \Rightarrow b \twoheadrightarrow d \text{ and } c \twoheadrightarrow d).$$

(iv) $\rightarrow$ is *confluent* or *Church–Rosser*, or has the *Church–Rosser property* (CR) if

$$\forall a, b, c \in A, \exists d \in A \qquad (a \twoheadrightarrow b \text{ and } a \twoheadrightarrow c \Rightarrow b \twoheadrightarrow d \;\&\; c \twoheadrightarrow d).$$

The definitions of WCR and CR are illustrated by Figs. 1a and b, respectively. (The dotted lines denote existential quantification.) Often the CR property is defined as suggested in Fig. 1c (and confluence as in 1b); but one easily proves that the two are equivalent. For some of the arguments in this paper it is better to think in terms of 1c.

1.2. Definition. Let $\mathscr{A} = \langle A, \rightarrow \rangle$ be an ARS.

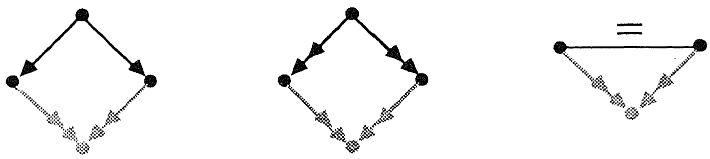
(i) We say that $a \in A$ is a *normal form* if there is no $b \in A$ such that $a \rightarrow b$. The set of normal forms of $\mathscr{A}$ is denoted by nf($\mathscr{A}$).

(ii) $\mathscr{A}$ (or $\rightarrow$) has the *unique normal form property with respect to reduction* ($UN^{\twoheadrightarrow}$) if $\forall a, b, c \in A$ ($a \twoheadrightarrow b$ and $a \twoheadrightarrow c$ and b, c are normal forms $\Rightarrow b \equiv c$).

(iii) $\mathscr{A}$ (or $\rightarrow$) has the *unique normal form property with respect to convertibility* ($UN^{=}$) if $\forall b, c \in A$ ($b = c$ and b, c are normal forms $\Rightarrow b \equiv c$). In conformance with most of the literature, we will henceforth denote the latter notion by UN, and refer to it as the "unique normal form property" without more.

(iv) $\mathscr{A}$(or $\rightarrow$) has the *normal form property* (NF) if $\forall a, b \in A$ (a is a normal form and $a = b \Rightarrow b \twoheadrightarrow a$).

The normal form property should not be confused with the property of *weak normalization* (WN), expressing that every element has a normal form, nor with the property of *strong normalization* (SN), expressing that every reduction sequence $a_0 \rightarrow a_1 \rightarrow a_2 \rightarrow \cdots$ must end, eventually, in a nor-



(a) (b) (c)

Figure 1

mal form. The reduction systems that we are primarily concerned with in this paper are neither SN nor WN, and sometimes even not CR. (They are all WCR, however.) The first implication of the following theorem is used only in 6.1; it is known as Newman's lemma.

1.3. THEOREM. (WCR *and* SN) $\Rightarrow$ CR $\Rightarrow$ NF $\Rightarrow$ UN $\Rightarrow$ $UN^{\twoheadrightarrow}$.

Proof. Easy. ∎

1.4. *Remark.* Note that UN and $UN^{\twoheadrightarrow}$ are equivalent in finite acyclic ARSs, but not in general. An example of an ARS satisfying the latter but not the former is the one in Fig. 2, consisting of five elements and reduction steps as displayed.

For the formulation of our results we need the notions of a consistent equational system and of a conservative extension. Also these notions can be defined already for ARSs:

1.5. DEFINITION. Let $\mathcal{A} = \langle A, \to_\alpha \rangle$ and $\mathcal{B} = \langle B, \to_\beta \rangle$ be two ARSs. Then $\mathcal{A}$ is a *sub-ARS* of $\mathcal{B}$, notation $\mathcal{A} \subseteq \mathcal{B}$, if:

(i) $A \subseteq B$

(ii) α is the restriction of β to A, i.e., $\forall a, a' \in A$ $(a \to_\beta a' \Leftrightarrow a \to_\alpha a')$

(iii) A is closed under β, i.e., $\forall a \in A$ $(a \to_\beta b \Rightarrow b \in A)$.

The ARS $\mathcal{B}$ is also called an *extension* of $\mathcal{A}$.

1.6. DEFINITION. (i) Let $\mathcal{A} < \langle A, \to \rangle$ be an ARS. Then $\mathcal{A}$ is *consistent* if not every pair of elements in A is convertible.

(ii) Further, let $\mathcal{A} = \langle A, \to_\alpha \rangle$, $\mathcal{B} = \langle B, \to_\beta \rangle$ be ARSs such that $\mathcal{A} \subseteq \mathcal{B}$. Then we define: $\mathcal{B}$ is a *conservative* extension of $\mathcal{A}$ if $\forall a, a' \in A$ $(a =_\beta a' \Leftrightarrow a =_\alpha a')$.

The proof of the following proposition is trivial.

1.7. PROPOSITION. (i) *If $\mathcal{A}$ is confluent and has two different normal forms, then $\mathcal{A}$ is consistent.*

(ii) *A conservative extension of a consistent ARS is again consistent.*

(iii) *A confluent extension $\mathcal{B}$ of $\mathcal{A}$ is conservative.*

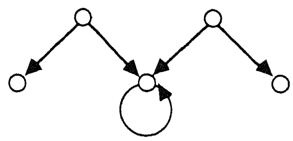


FIGURE 2

In the sequel we will deal with ARSs $\langle A, \rightarrow \rangle$, where A is a set of terms and the reduction relation $\rightarrow$ is generated by some reduction rules. All concepts introduced thus far (WCR, CR, UN, NF, consistency, conservativity) now apply to these reduction systems. Instead of "reduction systems" one can also adopt the phrase "term rewriting systems" or TRSs, although that name is usually reserved for cases where no bound variables are around. In Klop (1980) reduction systems of the kind we will consider in this paper are called "combinatory reduction systems" or CRSs.

2. Analyzing $\lambda\pi$: First Approach

We will suppose familiarity with the syntax of "pure" λ-calculus; that is, the equational system with terms built from variables by means of application and λ-abstraction and subject to the β-rule only. Likewise the corresponding reduction system, with the β-reduction rule, will be supposed known. As to the latter, we will use without further explanation terminology such as "redex" and "descendant." As a general reference, one may consult Barendregt (1981).

Let $\lambda\pi$ be the extension of the pure λ-calculus, or λ for short, with the constants π, π_0, and π_1 and with the following axioms (see Table I), which express that π, with the projections π_0 and π_1, is a surjective pairing. The set of (possibly open) $\lambda\pi$-terms will be denoted by $\Lambda\pi$, the set of pure λ-terms by Λ.

The equational system $\lambda\pi$ is our primary interest in this paper. The foremost problem that is posed now is to establish the *consistency* of $\lambda\pi$, that is, to show that not all terms of $\Lambda\pi$ are convertible to each other. As is well known, this concern is no luxury; some early axiomatizations of λ-calculus and extensions were inconsistent. Three methods to establish the consistency suggest themselves immediately:

I. *Showing that $\lambda\pi$ is a definitional extension of λ.* Unfortunately, this method is not applicable. It is easy to find λ-terms P, P_0, P_1 such that $P_0(PXY) = X$ and $P_1(PXY) = Y$ is derivable in λ for all $X, Y \in \Lambda$; so $\lambda\pi$ minus axiom SP ("λ-calculus + pairing") is a definitional extension of λ and hence consistent. However, in pure λ-calculus *surjective* pairing is not

TABLE I

$\lambda\pi$

Fst	$\pi_0(\pi XY) = X$
Snd	$\pi_1(\pi XY) = Y$
SP	$\pi(\pi_0 X)(\pi_1 X) = X$

definable. That is, there do not exist P, P_0, and P_1 in Λ, such that the equations Fst, Snd, and SP of Table I for P, P_0, and P_1 instead of π, π_0, and π_1, respectively, are derivable in λ. This result is due to Barendregt (1974). A short proof of the non-definability of $\lambda\pi$ in λ can also be found in Appendix 1 to Chapter 1 of de Vrijer (1987).

II. *Constructing a model for* $\lambda\pi$. There is a short and elegant model-theoretic proof of the consistency of $\lambda\pi$ via the *graph model* $P\omega$ of Plotkin and Scott. See, e.g., Scott (1975) or Exercise 18.5.12 in Barendregt (1981).

III. *Proving confluence (the Church–Rosser property) for the reduction system with terms* $\Lambda\pi$, *the* β*-reduction rule and the reduction rules in Table* II. It is clear that confluence of this reduction system would indeed entail the consistency (see Proposition 1.7(i)). The question whether confluence holds for this reduction system was posed in Mann (1973) (see also Barendregt, 1974; Böhm, 1975; Staples, 1975). Following de Vrijer (1987) we use the notation $\lambda\pi^c$ to refer to the system with this reduction relation; the ^{c} stands for "classical." (In Klop (1980) the system $\lambda\pi^c$ is called $\lambda + \mathrm{SP}$.) It seems to be taken for granted in most of the literature that π_0, π_1, and π^c are the natural reduction rules corresponding to the axioms for surjective pairing.

However, also this syntactic approach fails: $\lambda\pi^c$ is not confluent. In Klop (1980) the following counterexample is constructed. We use the $\lambda\pi$-terms

$$Y_{\mathrm{T}} \equiv (\lambda ab \cdot b(aab))(\lambda ab \cdot b(aab))$$

$$\Omega \equiv (\lambda x \cdot xx)(\lambda x \cdot xx)$$

$$C \equiv Y_{\mathrm{T}} \lambda cx \cdot \Omega[\pi(\pi_0 x)(\pi_1(cx))]$$

$$A \equiv Y_{\mathrm{T}} C.$$

Here Y_{T} is known as "Turing's fixed point combinator." The term Ω is to be perceived as an "inert symbol"; a variable x (or new constant) could play the same role. The "typical" reductions for these four terms are, respectively,

$$Y_{\mathrm{T}} M \twoheadrightarrow M(Y_{\mathrm{T}} M)$$

$$\Omega \rightarrow \Omega$$

$$CM \twoheadrightarrow \Omega[\pi(\pi_0 M)(\pi_1(CM))]$$

$$A \twoheadrightarrow CA$$

Let us furthermore introduce the abbreviations $\square \equiv CA$, $\square' \equiv \Omega\square$,

TABLE II

$\lambda\pi^c$

π_0	$\pi_0(\pi XY) \to X$
π_1	$\pi_1(\pi XY) \to Y$
π^c	$\pi(\pi_0 X)(\pi_1 X) \to X$

$\Box'' \equiv C\Box'$. By $M \downarrow N$ we denote that M and N have a common reduct, i.e., $\exists P\ M \twoheadrightarrow P$ and $N \twoheadrightarrow P$.

2.1. PROPOSITION. (i) $\Box \twoheadrightarrow \Box'$ *and* $\Box \twoheadrightarrow \Box''$,

(ii) *not* $\Box' \downarrow \Box''$.

Proof. (i) We have

$$A \twoheadrightarrow CA \twoheadrightarrow \Omega[\pi(\pi_0 A)(\pi_1 \Box)] \twoheadrightarrow \Omega[\pi(\pi_0 \Box)(\pi_1 \Box)] \to \Box'.$$

Deleting the first part of this reduction we have $CA \equiv \Box \twoheadrightarrow \Box'$. Furthermore, since $A \twoheadrightarrow \Box'$, also $CA \equiv \Box \twoheadrightarrow C\Box' \equiv \Box''$.

(ii) We give an intuitive argument. Note that

$$\begin{aligned} &\Box' \downarrow \Box'' &&\Leftrightarrow_{(1)} \\ &\Box' \downarrow \Omega[\pi(\pi_0 \Box')(\pi_1 \Box'')] &&\Leftrightarrow_{(2)} \\ &\Box \downarrow \pi(\pi_0 \Box')(\pi_1 \Box'') &&\Leftrightarrow_{(3)} \\ &\Box' \downarrow \Box''. \end{aligned}$$

Here $\Leftarrow_{(1)}$ is obvious. The other direction will not be treated here in full detail; a proof sketch is as follows: since $\Box'$ starts with the inert symbol Ω, and $\Box''$ does not, the best way to find a common reduct seems to be to perform the typical reduction

$$\Box'' \twoheadrightarrow \Omega[\pi(\pi_0 \Box')(\pi_1 \Box'')].$$

The equivalence $\Leftrightarrow_{(2)}$ holds since both $\Box'$ and $\Omega[\pi(\pi_0 \Box')(\pi_1 \Box'')]$ start with Ω, which, as it is inert, can be removed. As to $\Leftrightarrow_{(3)}$, this is also obvious, noting that no reduct of $\Box'$ starts with π as the first symbol.

So, in order to find a common reduct for $\Box$ and $\pi(\pi_0 \Box')(\pi_1 \Box'')$, the terms $\Box'$, $\Box''$ must be brought "in balance" to make the π^c-rule applicable; but that is the original problem. So the proof attempt is circular. ▮

The precise proof of part (ii) of the proposition can be found in Klop (1980). The following theorem summarizes the salient facts about $\lambda\pi^c$.

2.2. THEOREM. (i) *$\lambda\pi^c$ is not CR.*

(ii) *$\lambda\pi^c$ is not NF.*

Proof. (i) This is Proposition 2.1. For (ii) we use again the terms defined above. We further abbreviate $I \equiv \lambda x \cdot x$ and $\langle M \rangle \equiv \lambda x \cdot xM$; so $\langle M \rangle N \to NM$. Now

$$\pi(\pi_0\langle \Box \rangle)(\pi_1\langle \Box \rangle)(\lambda a \cdot I) \twoheadrightarrow I,$$

and also, using Proposition 2.1(i),

$$\pi(\pi_0\langle \Box \rangle)(\pi_1\langle \Box \rangle)(\lambda a \cdot I) \twoheadrightarrow \pi(\pi_0\langle \Box' \rangle)(\pi_1\langle \Box'' \rangle)(\lambda a \cdot I).$$

Due to the irreversible divergence of $\Box$ into $\Box'$ and $\Box''$, the last term however does not reduce to I. Hence NF does not hold. ∎

2.3. *Remark.* In Hardin (1986, 1987) a counterexample to NF for $\lambda\pi^c$ is given directly; a fortiori this is a counterexample to CR. Another advantage of Hardin's counterexample is that it does not depend on the standardization theorem, which was used in Klop (1980).

3. UNIQUE NORMAL FORMS FOR A RELATED SYSTEM

We saw in Section 2 that the system $\lambda\pi^c$ is not Church Rosser, and a closer analysis reveals that the main obstacle in an attempted CR proof results from the "non-left-linearity" of the rule π^c: the metavariable X occurs twice in the π^c-redex $\pi(\pi_0 X)(\pi_1 X)$, thus causing the redex to be unstable under reduction in one of the X's. (I.e., if $X \to X'$ then the redex $\pi(\pi_0 X)(\pi_1 X)$ ceases to be a redex after the reduction step $\pi(\pi_0 X)(\pi_1 X) \to \pi(\pi_0 X')(\pi_1 X)$.) Another complication lies in the ambiguity of the rules of $\lambda\pi^c$: the rules π_0 and π^c, and π_1 and π^c overlap. E.g., $\pi_0(\pi(\pi_0 X)(\pi_1 X))$ reduces to $\pi_0 X$ in two different ways: by applying rule π_0 on the whole term or by applying rule π^c. There are also some other types of overlap, which are easy to find.

As the factor of non-left-linearity was diagnosed to be the most serious one, it was proposed to isolate this phenomenon (by Hindley, see Böhm, 1975; or Staples, 1975), by studying the extension $\lambda\delta^h$ of λ which results from adding a single constant δ and the following simplified form of the π^c-rule:

$$\delta^h: \quad \delta XX \to X.$$

The system $\lambda\delta^h$ is investigated in Klop (1980). It was found there to be one of a few related systems which lack the Church Rosser property, but nevertheless satisfy UN. The counterexamples to the Church–Rosser theorem for those systems are all along the lines of the one for $\lambda\pi^c$ described in Section 2. In Bunder (1985) certain quite general conditions are formulated under which the extension of the λ-calculus with a rule of the form $\delta XX \to A$, where A is a λ-term possibly containing the metavariable X, lacks confluence.

The question of UN for $\lambda\pi^c$ remained open in Klop (1980). It can now be settled on the basis of a result from de Vrijer (1987), which allows the reasoning for UN in $\lambda\delta^h$ to be transferred to $\lambda\pi^c$. In this section we first present the proof of UN for $\lambda\delta^h$. The results for $\lambda\pi^c$ are covered in Section 4.

Let $\lambda\delta$ be the equational variant of $\lambda\delta^h$. That is, $\lambda\delta$ has the conversion rules:

$$\beta: \quad (\lambda x \cdot M)N = [x := N]M;$$

$$\delta: \quad \delta XX = X.$$

(Here $[x := N]$ is the usual substitution operator.) Then the reduction rule δ^h: $\delta XX \to X$ can be conceived of as a restricted form of the more liberal *conditional* reduction rule:

$$\delta^l: \quad \delta XY \to X \quad \text{if} \quad \lambda\delta \vdash X = Y.$$

(The superscript ^{l} stands for "left.") In contrast to δ^h, this rule is stable under reduction: a descendant of a δ^l-redex is still a δ^l-redex. It is easy to prove that the system $\lambda\delta^l$ (the rule δ^l in combination with β) does satisfy the Church–Rosser property, and hence also UN. Note by the way that the conversion relations generated by the one step reduction relation of $\lambda\delta^h$ and that of $\lambda\delta^l$ are the same, viz. the "=" of $\lambda\delta$.

3.1. *Remark.* In de Vrijer (1987) it is pointed out that the system $\lambda\delta^{lr}$ which is obtained by extending $\lambda\delta^l$ with the rule

$$\delta^r: \quad \delta XY \to Y \quad \text{if} \quad \lambda\delta \vdash X = Y,$$

is also CR. The reason is that under this further extension of reduction the convertibility relation that is generated remains the same: a common reduct of $\lambda\delta^{lr}$-convertible terms can be found already by using only β- and δ^l-reduction.

3.2. THEOREM. *$\lambda\delta^h$ satisfies UN.*

Proof. Since UN for $\lambda\delta^l$ is an immediate consequence of CR, it will do to show that the normal forms of $\lambda\delta^h$ and $\lambda\delta^l$ coincide. Now

$\mathrm{nf}(\lambda\delta^{\mathrm{l}}) \subseteq \mathrm{nf}(\lambda\delta^{\mathrm{h}})$ is an immediate consequence of the fact that the reduction rule δ^{h} is a mere restriction of δ^{l}.

For the converse inclusion assume $N \in \mathrm{nf}(\lambda\delta^{\mathrm{h}})$. We use induction on N to show that $N \in \mathrm{nf}(\lambda\delta^{\mathrm{l}})$. Suppose N does contain a δ^{l}-redex δXY. Then by the condition on rule δ^{l} we must have $\lambda\delta \vdash X = Y$. And consequently, by the induction hypothesis applied to X and Y and UN for $\lambda\delta^{\mathrm{l}}$, even $X \equiv Y$. This already contradicts the assumption that N was a $\lambda\delta^{\mathrm{h}}$-normal form. ▮

The above method does not work for $\lambda\pi^{\mathrm{c}}$ without further ado. It will be instructive to try this out by first devising a system $\lambda\pi^{\mathrm{l}}$ and then attempting to prove CR for it.

4. Analyzing $\lambda\pi$: Second Approach

The syntactic consistency proof of $\lambda\pi$ in de Vrijer (1987) makes use of a modification of the reduction relation of $\lambda\pi^{\mathrm{c}}$, bearing some resemblance to the system $\lambda\delta^{\mathrm{l}}$ above. The modified system is called $\lambda\pi^{\mathrm{lr}}$; to contrast it with $\lambda\pi^{\mathrm{c}}$ its one-step reduction is denoted by $>$ (with reflexive transitive closure $\geqslant$). Note that in the definition of the rules of $\lambda\pi^{\mathrm{lr}}$, the convertibility relation "$=$" of $\lambda\pi$, defined in Section 2, is assumed.

4.1. Definition. The set of terms of the system $\lambda\pi^{\mathrm{lr}}$ is $\Lambda\pi$; its one-step reduction relation $>$ is generated by the reduction rules given in table III.

Here "l" and "r" stand for "left" and "right." Again one readily verifies that the equivalence relation generated by $>$ coincides with the convertibility relation "$=$" of $\lambda\pi$. So there is no need to distinguish conversion in $\lambda\pi^{\mathrm{lr}}$ (or $\lambda\pi^{\mathrm{c}}$) from conversion in $\lambda\pi$. Note that the rules "l" and "r" both imply the rule $\pi^{\mathrm{c}}: \pi(\pi_0 X)(\pi_1 X) > X$.

The following definition is needed for stating the main result on $\lambda\pi^{\mathrm{lr}}$.

4.2. Definition. By "$\approx$" we denote the least equivalence relation on $\Lambda\pi$ satisfying the clause

$$X_0 = Y_0,\ X_1 = Y_1 \Rightarrow C[\pi X_0 X_1] \approx C[\pi Y_0 Y_1], \qquad \text{for all "contexts" } C[\ ].$$

TABLE III

$\lambda\pi^{\mathrm{lr}}$

β	$(\lambda x \cdot M)N > [x := N]M$
π_0	$\pi_0(\pi X_0 X_1) > X_0$
π_1	$\pi_1(\pi X_0 X_1) > X_1$
l	$\pi(\pi_0 X)Y > X$ if $\lambda\pi \vdash \pi_1 X = Y$
r	$\pi Y(\pi_1 X) > X$ if $\lambda\pi \vdash \pi_0 X = Y$

4.2.1. EXAMPLE. One has, e.g., $\pi_0(\pi((\lambda x \cdot x)y)z) \approx \pi_0(\pi yz)$, and $\lambda y \cdot \pi((\lambda x \cdot x)y)z \approx \lambda y \cdot \pi yz$, but not $(\lambda x \cdot \pi xz)y \approx \pi yz$.

In effect, $\approx$ disregards replacement of occurrences of subterms in the scope of a π by convertible ones. Since there are no π's there, on Λ the relation $\approx$ is just syntactic identity ($\equiv$).

Now in de Vrijer (1987) the Church–Rosser property for $\lambda\pi^{\mathrm{lr}}$ is established modulo $\approx$, that is, in the form of the following theorem. The proof is rather complicated and we will not go into any of its details here. Instead, we show at once how this theorem can be used for inferring the consistency of $\lambda\pi$ and the conservativity of $\lambda\pi$ over λ; and, moreover, for establishing UN, both for $\lambda\pi^{\mathrm{lr}}$ and $\lambda\pi^{\mathrm{c}}$.

4.3. THEOREM (CR/$\approx$). *If $\lambda\pi \vdash M = N$, then there exist $\approx$-equivalent Q_0 and Q_1, such that $M \geqslant Q_0$ and $N \geqslant Q_1$* (Fig. 3).

4.3.1. EXAMPLE. An instructive test case for this theorem is the counterexample to confluence of $\lambda\pi^{\mathrm{c}}$ in Section 2. See the terms $\Box'$, $\Box''$ there. Obviously, $\lambda\pi \vdash \Box' = \Box''$. Now there are indeed converging reductions as follows:

$$\Box' \equiv \Omega\Box \geqslant \Omega\Box'';$$

$$\Box'' \geqslant \Omega[\pi(\pi_0\Box')(\pi_1\Box'')] >_{\mathrm{r}} \Omega\Box''.$$

The first reduction is also possible in $\lambda\pi^{\mathrm{c}}$; the second is not, namely as regards the $>_{\mathrm{r}}$-step, which is justified in $\lambda\pi^{\mathrm{lr}}$ because $\lambda\pi \vdash \pi_0\Box' = \pi_0\Box''$. Remarkably, we find an "exact" common reduct $\Omega\Box''$ and not merely one modulo $\approx$. (Cf. question (iii) in Section 5.)

4.4. THEOREM (de Vrijer, 1987). (i) *$\lambda\pi$ is a conservative extension of λ, i.e., if $M, N \in \Lambda$, then $\lambda\pi \vdash M = N \Rightarrow \lambda \vdash M = N$.*

(ii) *$\lambda\pi$ is consistent.*

Proof. (i) Assume $\lambda\pi \vdash M = N$ for $M, N \in \Lambda$. Find Q_0 and Q_1 as indicated in the statement of Theorem 4.3. Then, as $\geqslant$-reduction cannot

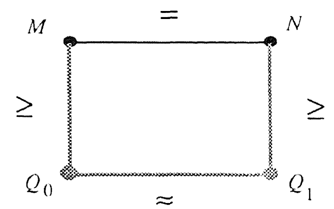


FIGURE 3

introduce constants which were not already present, all terms on the reduction sequences $M \geqslant Q_0$ and $N \geqslant Q_1$ must be in Λ, in particular $Q_0, Q_1 \in \Lambda$. Hence the reductions $M \geqslant Q_0$ and $N \geqslant Q_1$ are β-reductions and $Q_0 \equiv Q_1$. So M and N are convertible in λ as well.

(ii) Immediate by (i) and the consistency of λ (see Proposition 1.7(i)). ∎

4.5. THEOREM. *$\lambda\pi^{\text{lr}}$ satisfies NF and UN.*

Proof. By Theorem 1.3 it suffices to prove NF. We must verify that for normal forms N with $M = N$, one has also $M \geqslant N$. This will be accomplished by induction on the length of the normal form N. First notice that, since N is a normal form, the diagram in Fig. 3 here boils down to the diagram in Fig. 4.

Let P have the form

$$P \equiv \ldots \pi X_1 Y_1 \ldots \pi X_2 Y_2 \ldots \pi X_n Y_n \ldots,$$

with all maximal occurrences of the form πXY displayed ($n \geqslant 0$). Then, since $P \approx N$, there must be $X_i' = X_i$ and $Y_i' = Y_i$ (for $1 \leqslant i \leqslant n$) such that

$$N \equiv \ldots \pi X_1' Y_1' \ldots \pi X_2' Y_2' \ldots \pi X_n' Y_n' \ldots.$$

N is a normal form, hence so are the X_i' and the Y_i'. So the induction hypothesis yields $X_i \geqslant X_i'$ and $Y_i \geqslant Y_i'$ (for $1 \leqslant i \leqslant n$). Then by combining these reductions $P \geqslant N$ follows and, since we already had $M \geqslant P$, also $M \geqslant N$; this is the desired result. ∎

4.6. THEOREM. *$\lambda\pi^{\text{c}}$ satisfies UN.*

Proof. (The proof runs parallel to that of Theorem 3.2.) We show that the sets of $\lambda\pi^{\text{c}}$-normal forms and of $\lambda\pi^{\text{lr}}$-normal forms coincide: $\text{nf}(\lambda\pi^{\text{c}}) = \text{nf}(\lambda\pi^{\text{lr}})$. Then the result follows from Theorem 4.5 with the fact that the conversion relations of $\lambda\pi^{\text{c}}$ and $\lambda\pi^{\text{lr}}$ are the same. Of course, a $\lambda\pi^{\text{lr}}$-normal form is also a $\lambda\pi^{\text{c}}$-normal form, for the one step reduction relation of $\lambda\pi^{\text{c}}$ is a restriction of that of $\lambda\pi^{\text{lr}}$.

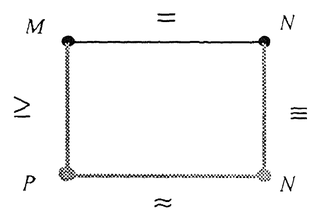


FIGURE 4

For the converse assume N to be a $\lambda\pi^{c}$-normal form. By induction on N we show that N cannot contain a $\lambda\pi^{lr}$-redex. Suppose it does. It must be an l- or an r-redex, say an l-redex $\pi(\pi_0 X)Y$. Note that X, being a subterm of N, is itself a $\lambda\pi^{c}$-normal form, too, and cannot be of the form $\pi X_0 X_1$; therefore also $\pi_1 X$ is a $\lambda\pi^{c}$-normal form, and it follows by the induction hypothesis that both $\pi_1 X$ and Y are $\lambda\pi^{lr}$-normal forms. Moreover, we have $\lambda\pi \vdash \pi_1 X = Y$, since the condition to $\pi(\pi_0 X)Y$ being an l-redex was supposed to be fulfilled. But then UN for $\lambda\pi^{lr}$ implies $\pi_1 X \equiv Y$, contradicting the assumption that N was a $\lambda\pi^{c}$-normal form. ∎

5. Assessment and Further Questions

The situation that is attained is summarized in Table IV. In the last column "cons" stands for "consistent and conservative." Here $\lambda(\eta)\pi^{\tau c}$ stands for typed λ-calculus (with or without η-reduction) extended with the rules and the corresponding constants, as in Table II. The results indicated by ⧺ are recent and were all derived from Theorem 4.3. (Consistency of $\lambda\pi^{lr}$ and $\lambda\pi^{c}$ was already known via the model theoretic method II of Section 2, but not conservativity of these systems over λ.) The NF and UN results indicated by ⧺ are new in this paper.

Some questions that remain and seem interesting enough to grant further research, are the following:

(i) The conservativity question for $\lambda\eta\pi$ (i.e., $\lambda\pi$ extended with the η-axiom) over $\lambda\eta$.

(ii) UN for $\lambda\eta\pi^{c}$ (i.e., $\lambda\pi^{c}$ extended with the η-reduction rule).

(iii) CR for $\lambda\pi^{lr}$.

(iv) Is there a system $(\Lambda\pi, \rightarrow)$ such that $\rightarrow$ is a restriction of the one step reduction relation $>$ of $\lambda\pi^{lr}$ satisfying the conditions (a)–(c) (and possibly (d)?

(a) $\rightarrow$ is decidable,

(b) $\rightarrow$ generates the convertibility relation of $\lambda\pi$,

TABLE IV

	WN	SN	CR	NF	UN	WCR	cons
$\lambda(\eta)\pi^{\tau c}$	+	+	+	+	+	+	+
λ	−	−	+	+	+	+	+
$\lambda\delta^{h}$	−	−	−	−	+	+	+
$\lambda\pi^{lr}$	−	−	?	⧺	⧺	?	⧺
$\lambda\pi^{c}$	−	−	−	−	⧺	+	‖

(c) $\rightarrow$ satisfies NF,

(d) $\rightarrow$ has the same normal forms as $\lambda\pi^{\text{lr}}$ (and hence $\lambda\pi^{\text{c}}$).

Or, a related question:

(v) Does an effective normal form strategy exist for $\lambda\pi^{\text{lr}}$?

Ad (i), (ii). η-conversion was not considered in de Vrijer (1987); it is not a priori clear whether the methods used there can be extended to cover η-reduction as well.

Ad (iii). The weaker CR/$\approx$ suffices for establishing the consistency/conservativity and UN results that are indicated by $\#$ in the table. These applications indicate that the reduction relation $>$ of $\lambda\pi^{\text{lr}}$ has at least proof-theoretical significance. Whether it can be considered as a sound computational concept remains doubtful, however. A positive answer to the CR question would shed some new light on this matter. We know of no reason why $>$ would not be CR (cf. Example 4.3.1).

Ad (iv). This question touches on a second aspect of $\lambda\pi^{\text{lr}}$ making it suspect as a reduction system that is natural from a computational point of view: its one-step reduction $>$ is not decidable. This follows from the undecidability of conversion in the pure λ-calculus; for $X, Y \in \Lambda$ we have

$$\pi(\pi_0 X)(\pi_1 Y) > X \Leftrightarrow \lambda \vdash X = Y.$$

Notice that one-step reduction in $\lambda\pi^{\text{c}}$ is decidable all right; but for $\lambda\pi^{\text{c}}$ one has the failure of CR and even of NF.

Ad (v). The existence of an effective normal form strategy would compensate for the lack of effectiveness of $>$.

6. Positive Results for Some Related Reduction Systems

As we have seen, obtaining confluence is highly problematic for reduction systems corresponding to $\lambda\pi$. We will now give a short survey of some positive confluence and unique normal form results for reduction systems which also have non-left-linear rewrite rules or rules related to the ones we have considered. As it turns out, certain more restrictive variants of $\lambda\pi$ (or $\lambda\delta$) yield a better chance to get confluence.

6.1. In $\lambda(\eta)\,\pi^{\tau c}$ (see Table IV) there is the restriction imposed by type constraints. Since the typed systems are strongly normalizing and WCR is easily checked, confluence is a consequence of Theorem 1.3 (see, e.g., Pottinger, 1981).

6.2. Let CL (combinatory logic) be the TRS with constants I, K, S and rules as in Table V. Furthermore, $\text{CL}\pi$, $\text{CL}\pi^{\text{c}}$, $\text{CL}\pi^{\text{lr}}$, $\text{CL}\delta$, $\text{CL}\delta^{\text{h}}$ are

TABLE V

CL

$SXYZ$	$\to XZ(YZ)$
KXY	$\to X$
IX	$\to X$

extensions analogous to $\lambda\pi$, etc. Like in the case of λ, the reduction systems $CL\pi^c$ and $CL\delta^h$ are not confluent (see Klop, 1980). We expect that the results of Section 4 hold also for the systems based on CL instead of λ.

Now suppose that $CL\pi^c$ is restricted by requiring that π_0, π_1 are unary operators and π is a binary operator. This means that π_0, π_1 always have an argument and that π always has two arguments. (In $CL\pi^c$ these three operators can be thought of as having "variable arity.") Call this restriction $CL\pi_f^c$. Confluence of $CL\pi_f^c$ is an immediate consequence of the result of Toyama (1987) that confluence of TRSs is preserved under disjoint sums. For λ an analogous statement holds, but then the extra restriction must be made that the arguments of the three operators are moreover closed terms (see Klop, 1980). Similar facts hold for δ, δ^h instead of π, π^c, respectively.

6.3. In Chew (1981) it is shown that TRSs including non-left-linear reduction rules have unique normal forms, provided the left-hand sides of the rules satisfy a suitable "non-overlapping" property. A corollary is that $CL\delta^h$ has the UN property; this result can also be obtained by a proof analogous to the one we gave for $\lambda\delta^h$ (Theorem 3.2).

A second application of Chew's theorem is the unique normal form property for CL plus applicative "parallel if," that is, with three extra constants C, T, and F (for "conditional," "true," and "false," respectively) and rules

$$CTXY \to X$$

$$CFXY \to Y$$

$$CZXX \to X.$$

It is explained in Chew (1981) that this case is essentially more complicated than the former one; the comparatively simple method we used for $CL\delta^h$ would not work now.

The system CL plus applicative "parallel if" is not confluent (Klop, 1980). In contrast it should be noted that the confluence of CL plus *ternary* "parallel if," with rules

$$\textit{if}\ T\ \textit{then}\ X\ \textit{else}\ Y \to X$$

$$\textit{if}\ F\ \textit{then}\ X\ \textit{else}\ Y \to Y$$

$$\textit{if}\ Z\ \textit{then}\ X\ \textit{else}\ X \to X,$$

follows by the main theorem of Toyama (1987) again—compare the case of $CL\pi_f^c$ mentioned in 6.2 above.

6.4. λ-calculus with *Church's δ-reduction* is the extension of the reduction system λ with the reduction rules

$$\delta XX \to \mathbf{0} \quad \text{if } X \text{ is a closed } \beta\delta\text{-normal form}$$

$$\delta XY \to \mathbf{1} \quad \text{if } X, Y \text{ are closed } \beta\delta\text{-normal forms and } X \not\equiv Y.$$

Here $\mathbf{0}$, $\mathbf{1}$ can be taken, e.g., as $\lambda xy \cdot x$ and $\lambda xy \cdot y$, respectively. A $\beta\delta$-normal form is a term without β-redexes and without subterms of the form δXY. In Mitschke (1977) this reduction system is shown to be confluent. A proof and generalizations can also be found in Barendregt (1981) and Klop (1980).

6.5. (To answer a question of J. P. Seldin, personal communication.) Related to the system with Church's δ-rules is the variant of $\lambda\pi^c$ which one gets by restricting the π^c-rule $\pi(\pi_0 X)(\pi_1 X) \to X$ to cases where X is a closed $\lambda\pi^c$-normal form. It is not hard to prove that β-reduction commutes with the π_0-, π_1-, and restricted π^c-rules. By the Lemma of Hindley and Rosen (see, e.g., Barendregt, 1981) confluence of the whole system then follows from confluence of β-reduction and confluence of the reduction relation generated by the other rules.

RECEIVED September 17, 1987; ACCEPTED January 30, 1988

REFERENCES

BARENDREGT, H. P. (1974). Pairing without conventional restraints, *Z. Math. Logik Grundlag. Math.* **20** 289–306.

BARENDREGT, H. P. (1981), The Lambda Calulus, Its Syntax and Semantics, 1st ed., North-Holland, 1981; 2nd ed., North-Holland, 1984.

BÖHM, C. (Ed.) (1975), "λ-Calculus and Computer Science Theory," Lecture Notes in Computer Science Vol. 37, Springer-Verlag, New York/Berlin.

BUNDER, M. (1985), An extension of Klop's counterexample to the Church-Rosser property to λ-calculus with other ordered pair combinators, *Theoret. Comput. Sci.* **39**, 337–342.

CHEW, P. (1981), Unique normal forms in term rewriting systems with repeated variables, *in* "13th Annual ACM Symposium on the Theory of Computing," pp. 7–18.

COUSINEAU, G., CURIEN, P.-L., AND MAUNY, M. (1985), The categorical abstract machine, *in* "Proceedings, Functional Programming Languages and Computer Architecture, Nancy, 1985" (J.-P. Jouannaud, Ed.), pp. 50–64, Lecture Notes in Computer Science Vol. 201, Springer-Verlag, New York/Berlin.

CURIEN, P.-L. (1986), Categorical combinators, sequential algorithms and functional programming, Research Notes in Theoretical Computer Science, Pitman, London.

HARDIN, TH. (1986), Yet another counter-example to confluence in λ-calculus with couples, preprint, INRIA, Rocquencourt.

HARDIN, TH. (1987), "Résultats de confluence pour les règles fortes de la logique combinatoire catégorique et liens avec les lambda-calculs," thèse de doctorat, Université Paris VII.

LAMBEK, J., AND SCOTT, P. J. (1986), "Introduction to Higher Order Categorical Logic," Cambridge Univ. Press, London.

KLOP, J. W. (1980), "Combinatory Reduction Systems." Mathematical Centre Tracts Vol. 127, Centre for Mathematics and Computer Science, Amsterdam.

MANN, C. R. (1973), "Connections between Proof Theory and Category Theory," dissertation, Oxford University.

MITSCHKE, G. (1977), "Discriminators in Combinatory Logic and λ-Calculus," Preprint 336, Technische Hochschule Darmstadt.

POTTINGER, G. (1981), The Church–Rosser theorem for typed λ-calculus with surjective pairing, *Notre Dame J. Formal Logic* **22**, 264–268.

SCOTT, D. (1975), Lambda calculus and recursion theory, *in* "Proceedings, Third Scandinavian Logic Symposium" (S. Kanger, Ed.), pp. 154–193, Amsterdam.

STAPLES, J. (1975), Church–Rosser theorems for replacement systems, *in* "Algebra and Logic" (J. N. Crossley, Ed.), Lecture Notes in Mathematics Vol. 450, pp. 291–307, Springer-Verlag, New York/Berlin.

TOYAMA, Y. (1987), On the Church–Rosser property for the direct sum of term rewriting systems, *J. Assoc. Comput. Mach.* **34**, No. 1, 128–143.

DE VRIJER, R. C. (1987), "Surjective Pairing and Strong Normalization: Two Themes in Lambda Calculus," dissertation, University of Amsterdam.