

STANDARD AND NORMAL REDUCTIONS

BY

R. HINDLEY

ABSTRACT. Curry and Feys' original standardization proof for $\lambda\beta$ -reduction is analyzed and generalized to $\lambda\beta\eta$ -reductions with extra operators.

There seem to be two slightly different definitions of 'standard reduction' in current use, without any awareness that they are different; it is proved that although these definitions turn out to be equivalent for $\lambda\beta$ -reduction, they become different for $\lambda\beta\eta$ and for reductions involving extra operators, for example the recursion operator.

Normal reductions are also studied, and it is shown that the basic normal-reduction theorem stays true when fairly simple operators like Church's δ and Curry's iterator Z are added, but fails for more complicated ones like the recursion operator R .

Finally, a table is given summarizing the results, and showing how far the main theorems on $\lambda\beta$ -reductions extend to reductions with various extra operators.

1. Introduction. Curry and Feys' standardization proof for the λ -calculus in [4, pp. 140–142] is not quite complete in itself, because it assumes implicitly that a certain reduction of residuals cannot be infinite. The present paper will set out in detail the implicit assumptions in Curry and Feys' proof (which is really an abstract proof and does not just apply to the λ -calculus) and show that these remain true when certain extra operators are added to the $\lambda\beta$ -calculus. (The finiteness of reductions of residuals has already been discussed in [9]; it was originally proved for the $\lambda\beta$ -calculus by D. E. Schroer in his thesis [15].)

Also, there seem to be two slightly different definitions of 'standard reduction' in current use; the original Curry-Feys definition, and a stronger definition, without any awareness that they are different. It will be shown here that these two definitions are equivalent for $\lambda\beta$ -reduction, but become distinct when certain operators are added, for example the recursion operator.

Normal reductions are closely related to standard reductions, so these will also be looked at here. The basic normal-reduction theorem has been generalized by Curry in [3] to reductions with certain 'generalized δ -operators'

Received by the editors August 30, 1976 and, in revised form, January 11, 1977 and March 10, 1977.

AMS (MOS) subject classifications (1970). Primary 02C20.

Key words and phrases. Reduction, lambda-conversion, combinator, standard reduction, normal reduction, Gross reduction, delta-conversion, recursion combinator.

© American Mathematical Society 1978

(including the iterator from which recursion can be defined); but it will be shown here that the theorem fails for the recursion operator itself. This failure, like the distinctness of the two definitions of standardness, will be traced to the failure of a certain lemma (Lemma 4 below).

Finally, the results of this paper and its two predecessors [8] and [9] will be set out in a table.

This paper was prepared in consultation with H. B. Curry, whom I wish to thank for very useful discussions, encouragement and advice. I also wish to record thanks to the Pennsylvania State University for kind hospitality and financial support during part of the work.

2. Notation and basic concepts. To save space, the notation of [9] will be used here without further explanation. The definitions are all in [9, §§2 and 5] except for the full definition of residuals, which is in [8, §§6 and 9].

However, one extra concept will be needed here, that of *seniority*. If R and S are two components of a term X , R is said to be *senior* to S iff either R lies entirely to the left of S , or R contains S as a proper part. Seniority is a nonreflexive linear ordering of the occurrences of terms in X , because R is senior to S iff the left-hand-most parenthesis of R is on the left of the left-most parenthesis of S .

DEFINITION 1 (FROM [4, p. 140]). *A standard reduction is a finite or infinite sequence $A_0 > A_1 > A_2 > \dots$ of contractions of redexes (ξ_0 in A_0 , ξ_1 in A_1 , ...) such that for each i , the redex ξ_i in A_i is not a residual of a redex in A_{i-1} that is senior to ξ_{i-1} .*

The purpose of calling certain reductions 'standard' is to pick out reductions whose structure is particularly simple, and then to prove that every finite reduction is equivalent, at least weakly, to one of these simple ones. We can then limit our study of reductions to 'standard' ones.

This has already been done for the $\lambda\beta$ -calculus. Curry and Feys in [4, pp. 140–142] have proved that every finite reduction is weakly equivalent to a standard one as defined above, and the structure of such standard reductions is indeed quite simple. To see the latter, note first that any λ -term X can be expressed uniquely as $X = \lambda x_1 \dots x_m \cdot (UV_1 \dots V_n)$, where $m \geq 0$, $n \geq 0$, and U is either an atom, or (only if $n \geq 1$) a term of form $(\lambda y \cdot U_1)$. If $n \geq 1$ and U is $\lambda y \cdot U_1$, then the redex $(\lambda y \cdot U_1)V_1$ in X is called the *head redex* of X , and the contraction of this redex is called a *head-contraction*. Then it is an easy exercise to see that any standard $\lambda\beta$ -reduction of X must have form

$$\begin{aligned} X &\gg Y && \text{by head-contractions,} \\ &= \lambda y_1 \dots y_p \cdot WZ_1 \dots Z_q && (W \text{ being an atom or } \lambda y \cdot W_1), \\ &\gg \lambda y_1 \dots y_p \cdot W'Z'_1 \dots Z'_q && \text{by standard reductions in } W, Z_1, \dots, Z_q. \end{aligned}$$

The reductions in W, Z_1, \dots, Z_q are called *internal reductions*.

For combinatory weak reduction, standard reductions are even simpler. Every combinatory term can be expressed uniquely as

$$X = cV_1 \dots V_n \quad (c \text{ an atom}),$$

and standard reductions have form

$$\begin{aligned} X &\gg Y && \text{by head-contractions,} \\ &= dZ_1 \dots Z_q && (d \text{ an atom}), \\ &\gg dZ'_1 \dots Z'_q && \text{by internal standard reductions.} \end{aligned}$$

Now suppose we add extra axiom-schemes (a) to combinatory or $\lambda\beta$ -reduction. (See [9, §5].) The above definition of 'standard' will remain meaningful and we shall see later that every finite reduction is still equivalent to a standard one. But the structure of standard reductions, though still simpler than the nastiest arbitrary reductions, will not be as simple as above. For example, with the recursion operator **R**, the following reduction is standard:

$$X = \mathbf{R}xy((\lambda z \cdot \mathbf{0})z) > \mathbf{R}xy\mathbf{0} > x.$$

This reduction begins with an internal contraction and finishes with a head-contraction, but it is standard because the head-redex $\mathbf{R}xy\mathbf{0}$ is not a residual of any redex in X . (A similar example can also be constructed using the iterator, **Z**.)

The difference between this and pure combinatory or $\lambda\beta$ -reduction is that in any pure reduction $X > Y > Z$ in which $X > Y$ is an internal contraction, the head-redex in Y must be the residual of a head-redex already in X ; internal contractions cannot create new head-redexes.

However, despite the extra complexity of standard reductions with (a)-redexes, the definition might have some uses, so I shall include (a)-redexes in the standardization theorem below. (For particular axiom-schemes (a) a different definition of 'standard' might give a more useful theorem, but the best definition to use would probably depend very much on the particular structure of the axiom-scheme.)

Incidentally, combinatory strong reduction also has its standardness concept and standardization theorem (see [5, pp. 119–121]). It also has a normal-reduction theorem [5, pp. 96, 124]. But these are so different from the ones here (for example there is no obvious concept of 'residual') that I have made no attempt to include strong reduction in this paper.

3. The standardization theorem.

THEOREM 1. *Every finite reduction ρ is strongly equivalent to a standard reduction, in each of the following systems:*

- (i) $\lambda\beta$ -(a)-reductions satisfying (D1)–(D7) of [9, §5];

(ii) *combinatory weak reductions, with or without extra (a)-schemes satisfying (D1)–(D7);*

(iii) *$\lambda\beta\eta$ -(a)-reductions satisfying (D1)–(D9) of [9, §5].*

REMARK 1. Strong equivalence [9, §2] means that the two reductions have the same residuals of any redex. Past standardization theorems have only claimed weak equivalence, but their proofs in most cases in fact give strong equivalence.

REMARK 2. For pure $\beta\eta$ -reductions, Curry and Feys' original standardness definition is different from that used here. In a $\beta\eta$ -reduction the η -steps can always be pushed down to the end. (See [4, p. 132, Theorem 3¹].) So Curry and Feys defined a standard $\beta\eta$ -reduction to be a standard β -reduction followed by a standard η -reduction [4, p. 142]. Such standard reductions are a bit simpler in structure than reductions in which β - and η -steps are mixed up together, so the Curry-Feys definition is better than Definition 1 above in this case.

The standardization theorem for Curry and Feys' definition follows immediately from the postponement of η -contractions, standardization for pure β , and standardization for pure η (which is easily proved by induction on the lengths of terms, since η -contractions shorten terms).

However, when (a)-redexes are added to the $\beta\eta$ -system the η -postponement theorem fails. For example, consider the reduction

$$\begin{aligned} X &= \mathbf{Z}(\lambda x \cdot 1x)uv \quad \text{where } \mathbf{Z} \text{ is Curry's iterator} \\ &> \mathbf{Z}1uv \quad \text{by } \eta \\ &> (\lambda xy \cdot xy)uv \quad \text{by the } \mathbf{Z}\text{-axiom (see e.g. [9, §5, Example 2])} \\ &\gg uv \quad \text{by } \beta. \end{aligned}$$

So when extra redexes are present, Theorem 1 (iii) is the only known standardization result.

REMARK 3. The standardization theorem's first published proof, for $\lambda\beta$ -reduction, was in Curry and Feys' [4, pp. 140–142]. For combinatory weak reduction a comparatively simple proof was given by Curry in [5, pp. 32–33]. In [3], the $\lambda\beta$ -theorem was extended to systems with extra reduction-axioms satisfying some abstract conditions, which included the iterator \mathbf{Z} and Curry's generalized δ -reductions (see [9, §5, Examples 2 and 3]), but do not include

¹Theorem 3 of [4, p. 132] says that every $\beta\eta$ -reduction is weakly equivalent to one in which all the η -steps occur at the end. There is in fact an error in the proof of this theorem in [4], but this has been corrected by R. P. Nederpelt in his doctoral thesis [13]. Another proof is in [1, Chapter II, §6]. Note that this η postponement theorem only claims weak equivalence, not strong. I do not know if it holds for strong equivalence.

the recursion operator.^{1a} Theorem 1 above does include the recursion and other similar operators.

For pure $\lambda\beta$ -reduction, the simplest proof is a recent one by G. Mitschke in [12], which uses the same general method as Curry's combinatory proof. Both these proofs rely on the fact that internal contractions cannot create new head-redexes. Since this property fails when (α)-redexes are added, to prove Theorem 1 we must fall back on Curry and Feys' original method.

4. Proof of the standardization theorem. Induction on the length of ρ will be used. The basis ($\rho = 0$, an empty reduction) is trivial, because empty reductions are counted as standard. The induction-step is equivalent to the following lemma (from [4, p. 140], with a strong equivalence clause added).

LEMMA 1. *In any of the three systems of Theorem 1, let $\xi + \rho$ be a contraction followed by a standard reduction; then there is a standard reduction strongly equivalent to $\xi + \rho$.*

The proof of Lemma 1 below will be essentially just Curry and Feys' original proof from [4, pp. 140–141]. Curry and Feys wrote their proof in λ -calculus language, but it is really in fact an abstract proof, and does not use any intimate structure of λ -redexes at all. It could have been carried out in an abstract system, with points instead of terms, and directed lines from point to point instead of contractions. Of course, to do this one would have to assume that to each pair ξ, η of coinitial lines (lines starting at the same point) there corresponds a set ξ/η of lines (residuals of ξ) all starting at the end of η , and similarly η/ξ . A reduction would be defined as a directed series of lines, and a development as a reduction of residuals of some coinitial set. (All this has in fact been done in [6, §1] or [8, §2].)

The only other λ -calculus properties used by Curry and Feys are the following:

(1) *For each point, the lines starting at that point have a nonreflexive linear ordering called seniority, and each set of coinitial lines has a seniormost member.*

(2) *If α is a set of coinitial lines, and ξ_1, ξ_2 are lines senior to (and coinitial with) all the members of α , then any development of α will leave at most one residual of each of ξ_1, ξ_2 , and if ξ_1 is senior to ξ_2 , the residual of ξ_1 will be senior to that of ξ_2 (if both residuals exist).*

(3) *If ρ is a reduction from X to Y , then no line starting at Y can be a residual with respect to ρ of two lines starting at X .*

^{1a}The only one of the abstract conditions in [3] whose satisfaction by the above reductions is not immediately obvious is perhaps Curry's (E10). But that condition's satisfaction follows from the form of the standardization proof for $\lambda\beta$ in [4], or from the proof in the present paper.

(4) Property E^+ holds; that is, all complete developments of a cointial set are strongly equivalent.²

(5) For any cointial set α , the development obtained by first contracting the seniormost member of α , then the seniormost residual, and so on, must be finite.

To interpret the above assumptions in the λ -calculus, read 'term' for 'point', 'redex' for 'line', and 'set of redexes in a term' for 'set of cointial lines'.

These assumptions are satisfied by systems (i) and (ii) of the standardization theorem, and all but (4) are satisfied by (iii). (Assumption (1) follows from the finiteness of terms, (2) and (3) are easy to check, (4) is proved in [8], and (5) follows from the finiteness of all developments [9].)

The next step is to deduce Lemma 1 from (1)–(5). Although the deduction below is really abstract, it is written in 'term' and 'redex' notation to make it easier to read. First of all, define the *seniormost complete development (SCD)* of a set α of redexes in a term A to be the reduction obtained by contracting first the seniormost member of α , then the seniormost residual, and so on. (It is well defined by assumption (1) and finite by (5).)

LEMMA 2. Assuming (1)–(5), let α be a set of redexes in a term A , and let $A \gg B$ by an SCD, τ , of α . Let $B > C$ by a single contraction η . Then there is a standard reduction strongly equivalent to $\tau + \eta$.

PROOF OF LEMMA 2. Let τ have form $A = A_0 > A_1 > \dots > A_{m+1} = B$, where $A_i > A_{i+1}$ by contracting the seniormost residual of α . Let α_i be the set of all the residuals of α in A_i .

Suppose $\tau + \eta$ is not itself standard. Then η is a residual of a redex η_i in an A_i ($i \leq m$) with η_i senior to ξ_i . Choose the least such i . By assumption (3), η_i is uniquely defined. Let E be the result of contracting η_i (see Figure 1).

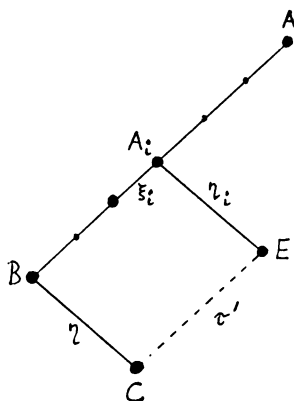


FIGURE 1

²The assumption of strong equivalence could not be replaced by weak equivalence here, not even if Lemma 1 was weakened to asserting only weak equivalence.

Let τ' be an SCD of the residuals of α in E . By assumption (4) applied to the set $\alpha_i \cup \{\eta_i\}$ in A_i , the end of τ' must be C ; also $\eta_i + \tau'$ is strongly equivalent to the reduction from A to C through B .

The reduction we want is $A \gg A_i > E \gg C$. It is strongly equivalent to $\tau + \eta$ by above. It is also standard, because $A \gg A_i > E$ is standard (by the choice of i as minimal), and $A_i > E \gg C$ is standard. This proves Lemma 2.

DEDUCTION OF LEMMA 1 FROM (1)–(5). Suppose ξ is a contraction from A to B , and ρ is a standard reduction

$$B = B_0 > B_1 > \dots > B_{n-1} > B_n = C,$$

where $B_i > B_{i+1}$ by contracting η_i . Apply Lemma 2 to A, B_0, B_1 . Then apply it to the resulting E, B_1 and B_2 . Continue in this way until C is reached (see Figure 2).

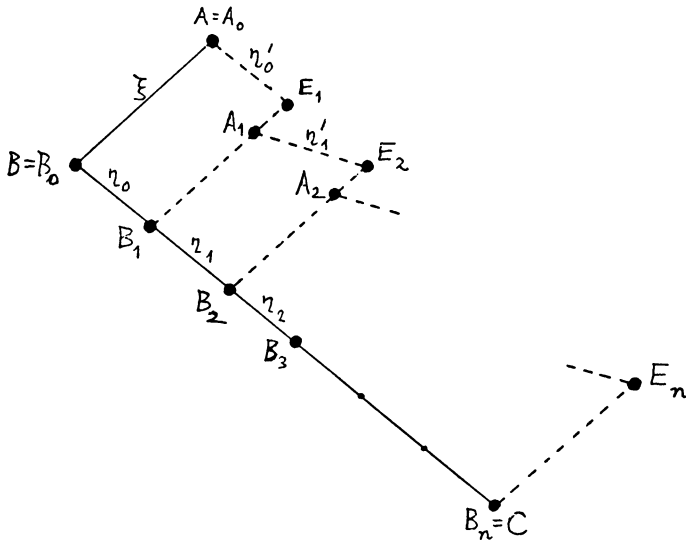


FIGURE 2

The result is a reduction

$$A = A_0 > E_1 \gg A_1 > E_2 \gg \dots \gg A_{n-1} > E_n \gg B_n = C,$$

where $A_i > E_{i+1}$ by contracting a redex η'_i whose residual in B_i is η_i , and $E_i \gg A_i \gg B_i$ by an SCD of the residuals of ξ in E_i .

The new reduction from A to C is strongly equivalent to the old one, by the strong equivalence in the proof of Lemma 2 applied n times.

The new reduction is also standard. Because it could only fail to be so if a section $E_i \gg A_i$ was null and η'_i was a residual of a redex ζ in A_{i-1} senior to η_{i-1} (see Figure 3).

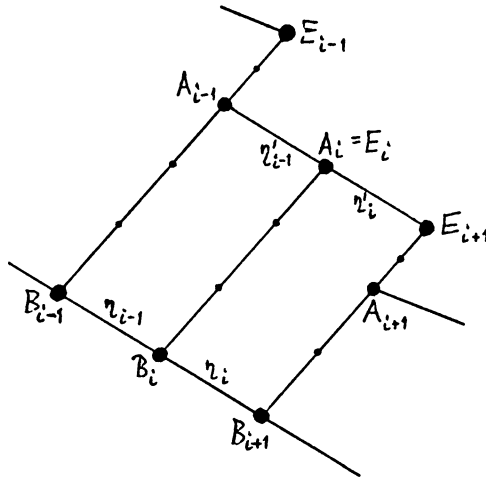


FIGURE 3

If this happened, A_{i-1} would contain (in order of seniority) ζ, η'_{i-1} , residuals of ξ . Hence by assumption (2), B_{i-1} would contain a residual ζ_{i-1} of ζ , senior to η_{i-1} . Then η_i would be the residual of this ζ_{i-1} , contrary to the standardness of the reduction from B_0 to B_n . (Because η_i would be a residual of ζ in A_{i-1} with respect to the reduction $A_{i-1} > A_i \gg B_i$, so by the strong equivalence in the proof of Lemma 2 it would be also a residual of ζ with respect to the other reduction, $A_{i-1} \gg B_{i-1} > B_i$; and hence by (3) it would be a residual of ζ_{i-1} in B_{i-1} .)

CONCRETE PROOF OF LEMMA 1. Lemma 1 has been deduced from (1)–(5), and we have seen that $\lambda\beta$ -(a)- and combinatory reductions satisfy (1)–(5). So now we only have to prove the lemma for $\lambda\beta\eta$ -(a)-reductions.

The latter reductions satisfy (1)–(3) and (5). If we could show that they satisfy (4) in the special situations where (4) was used in the preceding proof, then Lemma 1 would follow. But (4) was in fact only applied in the proof of Lemma 2, to the set $\alpha_i \cup \{\eta'_i\}$. In each case where Lemma 2 is used in the proof of Lemma 1, α_i is a set of residuals of ξ . Hence α_i cannot contain both β - and η -redexes at once. Also, η'_i has a residual η_i after the SCD of α_i .

Now any set α_i as above must be ‘well-separated’ ([8, §9] or [9, §2]). Then by Lemma 3 below, $\alpha_i \cup \{\eta'_i\}$ is also well-separated. But by [8, §9], all well-separated sets satisfy (4). So Lemma 1 is proved, given Lemma 3 below.

LEMMA 3. For $\lambda\beta\eta$ -(a)-reductions satisfying (D1)–(D9), if a term X contains a well-separated set α and a redex η , and some complete development ρ of α leaves at least one residual of η , then $\alpha \cup \{\eta\}$ is well-separated.

PROOF OF LEMMA 3. Suppose that on the contrary η was too close to a member ζ of α . By (4) applied to the well-separated set α , we could assume

that ζ was the first contraction in ρ . But contracting ζ would leave η without residuals, contrary to the conditions of the lemma.

5. The second definition of standardness. The following very natural standardness definition has appeared recently in unpublished notes, but the fact that it is not the same as the original definition seems to have been overlooked. We shall see below that the two definitions are equivalent for $\lambda\beta$ -reduction, but not for η -reduction or reduction with the recursion operator.

DEFINITION 2. Consider a finite reduction $\xi_0 + \xi_1 + \dots + \xi_n$, each ξ_i being a redex in a term A_i ; this reduction may be called standard (but will here be called *strongly standard*) iff for each i and each $j < i$, ξ_i is not a residual of a redex in A_j senior to ξ_j .

Recall that the original standardness definition only required ξ_i to be not a residual of a redex in A_j senior to ξ_j in the special case $j = i - 1$.

EXAMPLE 1. The following reduction, involving the recursion operator \mathbf{R} , is standard but not strongly standard. Let T be any redex whose contractum C is distinct from T , and let x, y be any variables.

$$\begin{aligned} X &= \mathbf{R}xT(\mathbf{R}(\sigma\mathbf{0})y\mathbf{0}) > \mathbf{R}xT(\sigma\mathbf{0}) > T\mathbf{0}(\mathbf{R}xT\mathbf{0}) \\ &> T\mathbf{0}(\mathbf{R}xC\mathbf{0}) = Y. \end{aligned}$$

This reduction is not strongly standard because the second T in the third stage is a residual of the T in the first stage, which is senior to the contracted redex. Moreover, there is no equivalent strongly standard reduction because any attempt to contract T earlier would produce not Y but a term with two C 's in it.

EXAMPLE 2. The following reduction involves only η -redexes, and it is standard but not strongly standard.

$$\begin{aligned} X &= \lambda x \cdot (\lambda v \cdot uv)(\lambda y \cdot xy) \\ &> \lambda x \cdot (\lambda v \cdot uv)x > \lambda v \cdot uv > u. \end{aligned}$$

But it is not such a strong example as Example 1, because there is an equivalent strongly standard reduction, which the reader can easily write out.

THEOREM 2. For pure $\lambda\beta$ -reduction or combinatory weak reduction, a reduction is strongly standard if and only if it is standard.

PROOF. Strong standardness obviously implies standardness. For the converse, note that by earlier remarks, a standard $\lambda\beta$ -reduction must have form

$$\begin{aligned} X &\gg U \quad \text{by head-contractions} \\ &= \lambda x_1 \dots x_p \cdot V_0 V_1 \dots V_q \quad (V_0 \text{ atomic if } q = 0) \\ &\gg \lambda x_1 \dots x_p \cdot V_0^* V_1^* \dots V_q^* \quad \text{by internal reductions} \\ &= Y. \end{aligned} \tag{6}$$

An easy induction on Y now gives the result.

THEOREM 3. *For pure η -reductions, i.e. with no β -redexes: although standardness is not equivalent to strong standardness, every reduction is still strongly equivalent to a strongly standard one.*

PROOF. Exercise.

REMARK. We have seen that for the β - and combinatory systems, strong standardness and standardness are equivalent. Although the two definitions are distinct for pure η -reduction, there is still a strong standardization theorem. But for \mathbf{R} -reductions strong standardization fails. So we cannot hope to extend it to arbitrary β -(a)-reductions. However, it can be extended to the system of intermediate generality, Curry and Feys' generalized δ -conversions ([4, p. 93] or [9, §5, Example 3]), as the next theorem will show. These systems include the iteration operator \mathbf{Z} and Church's δ (and from \mathbf{Z} one can define \mathbf{R}).

THEOREM 4. *For $\lambda\beta$ - or combinatory weak reduction with extra generalized δ -axioms satisfying the conditions of [4, p. 93] or [9, §5, Example 3], strong standardness is equivalent to standardness.*

PROOF. We cannot use the method of Theorem 2, because unfortunately standard reductions do not now have the simple form that they had before (in (6)).

Strong standardness obviously implies standardness. For the converse, consider any standard reduction

$$\rho = \eta_0 + \eta_1 + \eta_2 + \cdots + \eta_n,$$

where each η_i is a redex in a term A_i . If this reduction is not strongly standard, then there exist i , and $j < i$, such that η_i is a residual of a redex ξ in A_j senior to η_j . If $j = i - 1$, this would contradict the definition of standardness, so j must be less than $i - 1$. Consider the standard reduction $A_j \gg A_{i-1}$. If we could prove that ξ in A_j has a unique residual ξ' in A_{i-1} , senior to η_{i-1} , then η_i (being a residual of ξ) would have to be a residual of ξ' , contrary to the definition of standardness. This would prove the theorem.

The following lemma will prove the above required property of ξ .

LEMMA 4. *For the systems of Theorem 4, if $\eta_0 + \eta_1 + \cdots + \eta_n$ is a standard reduction, each η_i being a redex in a term X_i , and a redex ξ in X_0 is senior to η_0 , then ξ has exactly one residual in each X_i , and it is senior to η_i .*

PROOF OF LEMMA 4. By induction on n , it is enough to prove the case $n = 1$. Let $X_0 > X_1$ with ξ senior to η_0 in X_0 , and let η_1 in X_1 not be a residual of a redex in X_0 senior to η_0 . In X_0 , we have ξ properly containing η_0 or ξ completely to the left of η_0 . In X_1 , we have the contractum $C(\eta_0)$ of η_0 , and by the definition of residuals, ξ will have exactly one residual ξ' , either

properly containing $C(\eta_0)$ or lying completely to the left of $C(\eta_0)$.

Now η_1 in X_1 is either a 'new' redex (not a residual of anything at all) or else is a residual of a redex junior to η_0 . In the latter case, η_1 will be inside or to the right of $C(\eta_0)$ and hence must be junior to ξ' .

It remains to show that all new redexes in X_1 are junior to ξ' .

New redexes can be formed in two ways. First, the new redex could be formed inside the contractum of η_0 . But in this case, since ξ' is senior to $C(\eta_0)$, it will be senior to the new redex.

The second way is when the redex η_0 is part of a component T_0 of X_0 which becomes a redex T_1 when η_0 is contracted.

In this second case, if ξ is senior to T_0 in X_0 , then its residual ξ' will be senior to T_1 .

It remains to consider the case that ξ is not senior to T_0 . Since ξ is senior to η_0 which is in T_0 , in this case we must have T_0 contains ξ , ξ senior to η_0 . In fact T_0 must properly contain ξ , because if $T_0 = \xi$, then the redex T_1 would not be new at all, but the residual of ξ . Since T_0 properly contains ξ , we shall have in X_1 : T_1 properly contains ξ' , ξ' senior to $C(\eta_0)$. Hence the new redex T_1 cannot be a generalized δ -redex, because such redexes do not properly contain other redexes. (Cf. [9, §5, Example 3].)

The possibility remains that the new redex T_1 is a $\lambda\beta$ - or combinatory redex, say

$$T_1 = (\lambda x \cdot M)N \text{ or } SUVW \text{ or } K VW.$$

Now such a redex can only be created in one of the following ways:

- (i) $T_0 = \eta_0 N$, $C(\eta_0) = (\lambda x \cdot M)$, $T_1 = (\lambda x \cdot M)N$;
- (ii) $T_0 = \eta_0 W$, $C(\eta_0) = SUV$ or KV , $T_1 = SUVW$ or KVW ;
- (iii) $T_0 = \eta_0 VW$, $C(\eta_0) = SU$ or K , $T_1 = SUVW$ or KVW ;
- (iv) $T_0 = \eta_0 UVW$, $C(\eta_0) = S$, $T_1 = SUVW$.

Recall that the redex ξ is a proper part of T_0 and is also senior to η_0 . In cases (i) and (ii), this is impossible because η_0 is the seniormost proper part of T_0 . In cases (iii) and (iv), ξ would have to have form

$$\xi = \eta_0 V \text{ or } \eta_0 U \text{ or } \eta_0 UV, \tag{7}$$

and this is impossible. (ξ could not be a generalized δ -redex because such redexes contain no other redexes, and ξ could not be a β - or combinatory redex because no such redexes have the form (7).)

This proves Lemma 4, and hence completes the proof of Theorem 4.

REMARK 5. Lemma 4 is the key to the equivalence between standardness and strong standardness, and the key to Lemma 4 is the fact that new redexes cannot be senior to the residuals of old seniormost redexes. This property, and Lemma 4, fail for η -reductions and for \mathbf{R} -reductions. (See Examples 1 and 2, §5.) For instance, in Example 1 the seniormost redex T at the

beginning had its residual 'surrounded' by a newly created redex at the second stage of the reduction; the result was a reduction that was standard but not strongly standard.

REMARK 6. $\beta\eta$ -reduction. Is there a strong standardization theorem for the $\beta\eta$ -system? As noted above, strong standardness will be distinct from standardness whenever η -redexes are present, but the possibility of a strong standardization theorem has not been ruled out.

For pure $\beta\eta$ -reduction one can define a strongly standard reduction to be a strongly standard β -reduction followed by a strongly standard η -reduction. Then the η -postponement theorem implies that every reduction is at least weakly equivalent to a strongly standard one.

For $\beta\eta(a)$ or $\beta\eta\delta$ -reductions the postponement of η fails, but if we defined 'strongly standard' exactly as in Definition 2 without separating β - and η -redexes, a strong standardization theorem might hold.

6. Normal reductions. The main use of normal reductions is in proofs that terms do not have normal forms.³ One proves that if any reduction of X reaches a normal form, then the normal reduction will; so if the normal reduction of X is infinite, all reductions of X will be infinite.

Curry and Feys' original definition of normal reduction (for the $\lambda\beta$ -calculus) was a finite standard reduction whose end-term contains no redexes. (See [4, p. 140].) With this definition it follows at once from the standardization theorem that if X has a normal form X^* , then a normal reduction of X must exist and end at X^* [4, p. 142]. But this definition is not very easy to use. A much more natural one is Curry's later one, stated below, which has now been generally adopted.

DEFINITION 3. The *normal reduction* of a term X is the reduction obtained by first contracting the seniormost redex in X , then the seniormost redex in the result, and so on, continuing as long as possible. (The normal reduction may be infinite.)

THEOREM 5. For $\lambda\beta$ - or combinatory weak reduction with extra generalized δ -axioms satisfying the conditions of [4, p. 93] or [9, §5, Example 3], for finite reductions the above definition of normality is equivalent to Curry and Feys' original definition [4, p. 140].

PROOF. Consider a standard reduction which ends at a term containing no redexes. This reduction must satisfy the above normality definition, because if ξ was a seniormost redex which was not contracted, then by Lemma 4 there would be a residual of ξ in the end-term.

³A term X is said to have normal form Y iff X reduces to Y and there are no redexes in Y . The Church-Rosser theorem, which holds for all the systems in the present paper, implies uniqueness of normal forms.

THEOREM 6 (NORMAL-REDUCTION THEOREM). *For $\lambda\beta$ - or combinatory weak reductions with extra generalized δ -axioms satisfying the conditions of [4, p. 93] or [9, §5, Example 3], if a term X has a normal form X^* , then the normal reduction of X is finite and ends at X^* .*

PROOF. By Theorems 1 and 5. (This theorem was first proved in full by Curry at the end of [3]; but for $\lambda\beta$ -reduction an earlier proof existed in the 1973 manuscript of Barendregt's mentioned in the next section.)

REMARK 1. The normal-reduction theorem does not generalize from δ -reductions to arbitrary reductions satisfying (D1)–(D7). For example, given the recursion operator \mathbf{R} , consider the term

$$\mathbf{R}xT(\mathbf{R}0g0), \tag{8}$$

where x and g are variables and T is any term without a normal form. The term (8) has normal form x , but its normal reduction proceeds entirely inside T and is infinite.

The 'reason' for the failure of the normal-reduction theorem here is the same as for the failure of strong standardization; Lemma 4 fails.

Our only hope of getting a very general normal-reduction theorem is to change the definition of 'normal reduction'. The new definition would have to force $\mathbf{R}0g0$ to be contracted in (8), in between the contractions in T . One reduction with this property is the following, which I believe was originally suggested by W. Gross for a slightly different purpose.

DEFINITION 4. A *Gross reduction* of a term X consists of first a complete development of all the redexes in X , then a complete development of all the redexes in the result, and so on, as long as possible.⁴

THEOREM 7 (W. GROSS). *If $A \gg B$ by $\lambda\beta$ - or combinatory weak reductions augmented by (a)-schemes satisfying (D1)–(D7), then any Gross reduction of A will reach a B' such that $B \gg B'$.*

PROOF. Suppose we have a reduction $\xi_0 + \xi_1 + \dots + \xi_n$, with ξ_i in A_i , giving

$$A = A_0 > A_1 > A_2 > \dots > A_{n+1} = B. \tag{9}$$

Let a Gross reduction of A be

$$A = A'_0 \gg A'_1 \gg A'_2 \dots, \tag{10}$$

where $A'_j \gg A'_{j+1}$ by a complete development of all the redexes in A'_j . (By Curry's property (E), it does not matter which complete development we use.) We shall see by induction on n that $A_n \gg A'_n$. The basis is $n = 0$, and we have $A_0 = A'_0$. For the induction step, suppose $A_{n-1} \gg A'_{n-1}$. Let ξ_{n-1} be the redex

⁴I think the name 'Gross reduction' has also been given to another kind of reduction in the literature.

whose contraction changes A_{n-1} to A_n . (See Figure 4.)

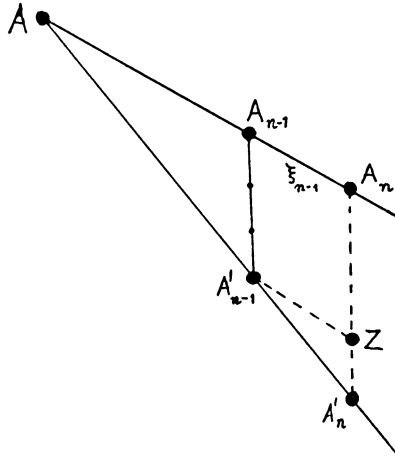


FIGURE 4

Now the lemma of parallel moves is true in the present situation. (See [4, p. 113, Theorem 5] or [8, §7].) Applying this lemma to A'_{n-1} , A_{n-1} and A_n gives us a Z such that

$$A_n \gg Z, \quad A'_{n-1} \gg Z,$$

and the reduction from A'_{n-1} to Z is a complete development of the residuals in A'_{n-1} of ξ_{n-1} . Let Z' be the result of performing a complete development of all the redexes in Z which are residuals of redexes in A'_{n-1} . Then the reduction $A'_{n-1} \gg Z \gg Z'$ is a complete development of all the redexes in A'_{n-1} , so Z' must be the same as A'_n , by (E). Hence $A_n \gg A'_n$, as required.

COROLLARY 7.1. *For $\lambda\beta$ - or combinatory weak reductions with extra (a)-schemes satisfying (D1)–(D7), if X has a normal form X^* , then the Gross reductions of X will end at X^* .*

REMARK 2. *Nearly normal reductions.* Proofs of nonexistence of normal forms would be made even easier if we had a larger class of reductions, called say ‘nearly normal reductions’, such that if any one of the nearly normal reductions of X was infinite, then all reductions of X would be infinite. One such class has been defined by H. P. Barendregt in an unpublished manuscript (1973); his definition is as follows.

DEFINITION 5. A *nearly normal reduction* of a term X is a reduction $X = X_0 \gg X_1 \gg X_2 \gg \dots$ such that for each i , $X_i \gg X_{i+1}$ either (i) by contracting the seniormost redex in X_i , or (ii) by reducing a component of X_i with a normal form to its normal form (provided the component is not itself already in normal form).

THEOREM 8 (BARENDREGT). *For $\lambda\beta$ - or combinatory weak reduction with extra generalized δ -axioms satisfying the conditions of [4, p. 93], if X has a normal form X^* , then all the nearly normal reductions of X end at X^* .*

PROOF. The theorem will follow immediately from the second of the following two lemmas. That lemma will be proved rather like Gross' Theorem above, but Gross' proof applied to arbitrary (a)-axioms, whereas the proof below will apply only to generalized δ -axioms. The ordinary normal-reduction theorem does not extend to arbitrary (a)-axioms, so of course Barendregt's stronger result cannot.

LEMMA 5. *For the systems in Theorem 8, if X contains a seniormost redex ξ , then ξ or its residual must be contracted in any nearly normal reduction of X that has more steps than there are components in X .*

PROOF. Let $c(X)$ be the number of components in X ; that is, $c(\text{atom}) = 1$, $c(YZ) = 1 + c(Y) + c(Z)$, and $c(\lambda v \cdot Y) = 1 + c(Y)$. By induction on $c(X)$, we shall see that in every nearly normal reduction with length $c(X) + 1$,

$$X = X_0 \gg X_1 \gg \cdots \gg X_{c(X)+1}, \quad (11)$$

the steps $X_i \gg X_{i+1}$ cannot all be of type (ii) in the definition of nearly normal reduction.

If X is an atom or $\lambda v \cdot Y$, the proof is trivial. Now suppose $X = YZ$. Suppose every step in (11) was of type (ii), namely the reduction of a part T_i of X_i to its normal form T_i^* in X_{i+1} ($T_i^* \neq T_i$). Then by the induction-hypothesis, (11) could not take place entirely in Y and Z . There would have to be a step whose T_i was the whole of X_i , and we would have $i \leq c(Y) + c(Z)$. Hence (11) would have form

$$X = X_0 \gg \cdots \gg X_i = T_i \gg T_i^*,$$

but then it could reduce no further, because T_i^* would be in normal form. This would be too short for (11).

Thus (11) must contain at least one (i)-step. This must contract the residual of ξ , if it has not already been contracted in an earlier (ii)-step.

LEMMA 6. *For the systems of Theorem 8, if X is any term and a finite or infinite nearly normal reduction of X is $X = X_0 \gg X_1 \gg X_2 \gg \cdots$, and the normal reduction of X is $X = Y_0 \gg Y_1 \gg Y_2 \gg \cdots$, then each Y_n reduces to some X_{j_n} .*

PROOF. Induction on n is used. (See Figure 5.) For $n = 0$, choose $j_0 = 0$.

For $n + 1$, assume that j_n has already been chosen; call it j . Let ξ be the seniormost redex in Y_n , so that the step from Y_n to Y_{n+1} is the contraction of ξ . By the standardization theorem, there is a standard reduction from Y_n to

X_j , and by Lemma 4, ξ has at most one residual in X_j with respect to this reduction.

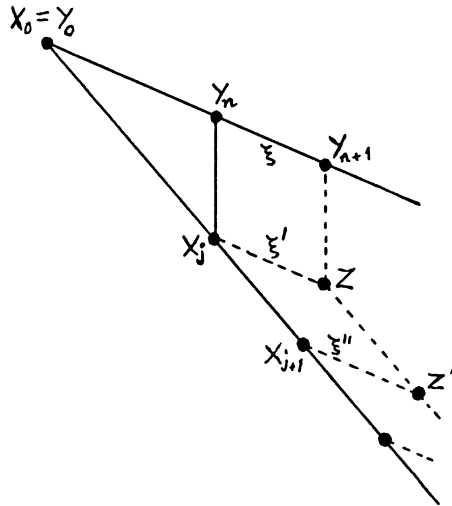


FIGURE 5

If ξ has no residual in X_j , choose $j_{n+1} = j$. Then the parallel-moves lemma [8, §7] gives $Y_{n+1} \gg X_j$.

If ξ has a residual ξ' in X_j , then the proof of Lemma 4 shows that ξ' is the seniormost redex in X_j . Let Z be the result of contracting ξ' in X_j . By the parallel-moves lemma, $Y_{n+1} \gg Z$. If Z is the next stage of the given nearly normal reduction, then we have $Y_{n+1} \gg X_{j+1}$, and we can choose j_{n+1} to be $j + 1$.

Now suppose Z is not the next stage of the given nearly normal reduction. Then the step from X_j to X_{j+1} must be of type (ii) in the definition of 'nearly normal', namely the reduction of some part T of X_j to its normal form T^* . By the normal-reduction theorem, there is a normal reduction from T to T^* . By Lemma 4, ξ' has at most one residual ξ'' in X_{j+1} with respect to this reduction, and as before, ξ'' is seniormost in X_{j+1} . By the parallel-moves lemma, there is a term Z' with $Y_{n+1} \gg Z'$, and either $X_{j+1} = Z'$, or $X_{j+1} > Z'$ by contracting ξ'' . (See the lower part of Figure 5.)

If $X_{j+1} = Z'$, choose $j_{n+1} = j + 1$. But if $X_{j+1} > Z'$, repeat the above argument again and again until a term Z^* is obtained with

$$Z^* = X_{j+k} \quad (\text{some } k).$$

This must happen eventually because Lemma 5 says that the steps $X_j \gg X_{j+1} \gg X_{j+2} \gg \dots$ cannot all be of type (ii). Choose $j_{n+1} = j + k$.

This completes Lemma 6 and hence Theorem 8.

REMARK 3. $\beta\eta$ -reductions. For pure $\beta\eta$ -reductions, a normal reduction is

usually taken to be a normal β -reduction followed (if the β -reduction is finite) by a normal η -reduction. The normal-reduction theorem for this definition is then a consequence of (i) the corresponding β -theorem, (ii) the postponement of η -contractions, (iii) the fact that every term has a normal form for pure η -reduction, and (iv) the fact that a term has a $\beta\eta$ -normal form if and only if it has a β -normal form ([5, Lemma 13.1, p. 124] or [1, Chapter II, Corollary 6.14]).

For $\beta\eta\delta$ -reductions, the η -postponement property fails, so in such systems a normal reduction must be defined by Definition 3 without distinguishing between β - and η -redexes. I do not know if there is a normal-reduction theorem for this definition.

For $\beta\eta$ -(a)-reductions in general even Definition 3 fails (at least, in the case of the recursion operator). But an analogue of Gross' Theorem probably holds.

7. Conclusion. In [7], [8], [9] and the present paper, the basic properties of λ -reduction have been extended to reductions with extra operators. Six main types of system have been considered:

β : $\lambda\beta$ - or combinatory weak reduction without extra axioms;

$\beta\delta$: $\lambda\beta$ - or combinatory weak reduction augmented by generalized δ -axioms satisfying the conditions of [4, p. 93] or [9, §5, Example 3] (including Church's δ and Curry's iterator \mathbf{Z});

β -(a): $\lambda\beta$ - or combinatory weak reduction augmented by (a)-schemes satisfying (D1)–(D7) (including the recursion operator \mathbf{R});

$\beta\eta$: $\lambda\beta\eta$ -reduction without extra axioms;

$\beta\eta\delta$: $\lambda\beta\eta$ -reduction with generalized δ -axioms;

$\beta\eta$ -(a): $\lambda\beta\eta$ -reduction with (a)-schemes satisfying (D1)–(D9).

The extension results can be summarized in the following table. Roughly speaking, $\beta\delta$ has all the properties of unmodified β -reduction, and β -(a) has all the properties that do not depend on Lemma 4.^{4a}

An interesting consequence of this is that if we wish to define the recursive functions when zero and successor are atomic constants added to the λ -

^{4a}Since some of these reduction-properties hold for $\beta\delta$ but fail for full β -(a), it is natural to ask where the borderline lies; are there other simple conditions like Curry's generalized δ which imply all the above properties?

The answer is 'yes'. For example, one quite neat condition is that all the extra axioms have form $av_1 \dots v_k > N$ where k depends on a , and N is a combination of constants and v_1, \dots, v_k . For such axioms, all the theorems that hold for β -(a) would hold as a special case. All the other theorems of this paper would also hold; because we saw that the key to their proofs was that contracting an internal redex does not create a new head redex, and this property holds for redexes of the above form.

However, if η -contractions were also allowed, then the Church-Rosser property could fail. (For example, consider $\lambda v_k \cdot av_1 \cdot \dots \cdot v_k$.)

system, we get good reduction-properties by adjoining Curry's iterator **Z** instead of the recursion combinator **R**. (**R** can be defined in terms of **Z**; [5, p. 223].)

For the $\beta\eta$ -systems, there are still some gaps in the table, for example strong standardization and the closely related normal-reduction properties.

Property	β	$\beta\delta$	$\beta(a)$	$\beta\eta$	$\beta\eta\delta$	$\beta\eta(a)$
Finiteness of developments	true	true	true	true	true	true
Curry and Feys' (E) and (E ⁺)	true	true	true	true for well-separated sets		
Parallel-moves lemma	true	true	true	true with slight change ⁵		
Church-Rosser theorem	true	true	true	true	true	true
Postponement of η -steps	—	—	—	true	false	false
Standardization theorem:						
{straightforward Definition 1	true	true	true	true	true	true
{ η -postponed definition	—	—	—	true	false	false
Standard = strongly standard:						
{straightforward	true	true	false for R	false	false	false
{ η -postponed	—	—	—	false	false	false
Strong standardization:						
{straightforward	true	true	false for R	?	?	false for R
{ η -postponed	—	—	—	true	false	false
Normal-reduction theorem:						
{straightforward	true	true	false for R	?	?	false for R
{ η -postponed	—	—	—	true	false	false
Nearly-normal reduction theorem	true	true	false for R	?	?	false for R
Gross' Theorem	true	true	true	true ⁶	?	?

REFERENCES

1. H. P. Barendregt, J. Bargstra, J. W. Klop and H. Volken, *Degrees, reductions and representability in the lambda calculus*, Math. Instituut, Budapestlaan 6, Utrecht, preprint 22, 1976.
2. A. Church and J. B. Rosser, *Some properties of conversion*, Trans. Amer. Math. Soc. **39** (1936), 472-482.
3. H. B. Curry, *A study of generalized standardization in combinatory logic*, Lecture Notes in Math., vol. 500, Springer-Verlag, Berlin and New York, 1974, pp. 45-55.
4. H. B. Curry and R. Feys, *Combinatory logic*. I, North-Holland, Amsterdam, 1958. MR **20** #817.
5. H. B. Curry, R. Hindley and J. P. Seldin, *Combinatory logic*. II, North-Holland, Amsterdam, 1972.

⁵See footnote to [8, §7].

⁶Proved in [1, Chapter II, §3].

6. R. Hindley, *An abstract form of the Church-Rosser theorem. I*, J. Symbolic Logic **34** (1969), 545–560.
7. ———, *An abstract Church-Rosser theorem. II*, J. Symbolic Logic **39** (1974), 1–21.
8. ———, *The equivalence of complete reductions*, Trans. Amer. Math. Soc. **229** (1977), 227–248.
9. ———, *Reductions of residuals are finite*, Trans. Amer. Math. Soc. **240** (1978), 345–361.
10. R. Hindley, B. Lercher and J. P. Seldin, *Introduction to combinatory logic*, Cambridge Univ. Press, London and New York, 1972. MR **49** #25.
11. J. M. E. Hyland, *A simple proof of the Church-Rosser theorem*, unpublished manuscript available from J. Hyland, Math. Inst., 24–29 St. Giles, Oxford, England.
12. G. Mitschke, *The standardization theorem for the λ -calculus*, 1975, unpublished manuscript available from G. Mitschke, Fachbereich Mathematik, Technische Hochschule, Darmstadt, Federal Republic of Germany.
13. R. P. Nederpelt, *Strong normalization in a typed lambda-calculus with lambda-structured types*, Doctoral thesis, 1973, Tech. Hogeschool, Eindhoven, The Netherlands.
14. J. B. Rosser, *A mathematical logic without variables*, Ann. of Math. **36** (1935), 127–150.
15. D. E. Schroer, *The Church-Rosser theorem*, Doctoral thesis, Cornell Univeristy, Ithaca, N. Y., 1965.

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY COLLEGE, SWANSEA, WALES, UNITED KINGDOM SA2 8PP