A General Church-Rosser Theorem

Peter Aczel

We prove the Church-Rosser theorem in a general framework. Our result easily yields the standard result for the lambda calculus but also has wider application.

§1. The Main Theorem

1.1 An expression system consists of an infinite set of variables and a set of forms. Each form has an arity i.e. a finite sequence $k_1, \ldots, k_m$ $(m \geq 0)$ of natural numbers. If $m = 0$ the form is a constant. If $k_1 = \ldots = k_m = 0$ the form is simple.

Expressions are inductively generated using the two rules:-

1) Every variable is an expression.

2) If F is a form with arity $k_1, \ldots, k_m$ $(m \geq 0)$ and $a_1, \ldots, a_m$ are expressions then $F((\vec{x}_1)a_1, \ldots, (\vec{x}_m)a_m)$ is an expression, where for $i = 1, \ldots, m$ $\vec{x}_i$ is a list of k_i variables.

The expression generated by 2) is said to have form F and parts $a_1, \ldots, a_m$. Free and bound occurrences of variables are defined in the usual way, so that occurrences of a variable in the list $\vec{x}_i$ that are free in a_i become bound in $F((\vec{x}_1)a_1, \ldots, (\vec{x}_m)a_m)$. Alphabetic variants of expressions are identified in the standard way.

Below we shall usually write $F(a_1, \ldots, a_m)$ instead of $F((\vec{x}_1)a_1, \ldots, (\vec{x}_m)a_m)$. It must be kept in mind that with this abuse of notation a variable that is free in a_i can become bound in $F(a_1, \ldots, a_m)$. Also, an expression $F(a_1, \ldots, a_m)$ may be the same as an expression $F(b_1, \ldots, b_m)$ while a_i is not the same as b_i for $i = 1, \ldots, m$.

1.2 We shall be concerned with a partial function on the expressions which we shall call a contraction operation. An expression in the domain of the operation will be called a redex and its value under the operation will be called the contraction of the redex. We shall insist that no variable is a redex. Each contraction operation generates a relation of definitional equality.

This is the smallest invariant equivalence relation such that each redex is definitionally equal to its contractum. Here a binary relation R is invariant if xRx for each variable x and $F(a_1, \ldots, a_m)RF(b_1, \ldots, b_m)$ whenever $a_1Rb_1, \ldots, a_mRb_m$ for each form F.

An expression is normal if no subexpression is a redex, where a variable has only itself as a subexpression and $F(a_1, \ldots, a_m)$ has itself and all the subexpressions of its parts $a_1, \ldots, a_m$ as subexpressions.

1.3 When is a contraction operation consistent in the sense that only identical expressions can be both normal and definitionally equal? The Church-Rosser property (CRP) is the key to answering this question. A binary relation R has the CRP if:-

aRb and aRc implies bRd and cRd for some d.

LEMMA

If the reduction relation has CRP then the contraction operation is consistent. The reduction relation is the smallest transitive invariant relation such that each redex reduces to its contractum.

PROOF. An expression is normal if it reduces only to itself. Hence consistency follows from:- a is definitionally equal to b if and only if a and b reduce to a common expression. This is easily seen to follow from the CRP for reduction.

1.4 The next lemma is easily proved and will be useful later.

LEMMA

(i) Every invariant relation is reflexive.

(ii) The transitive closure of an invariant relation is invariant.

(iii) The transitive closure of a relation with CRP also has CRP.

1.5 Let $>$ be the smallest invariant relation such that if $F(b_1, \ldots, b_m)$ is a redex with contractum b then

$$a_1 > b_1, \ldots, a_m > b_m \text{ implies } F(a_1, \ldots, a_m) > b.$$

LEMMA

Reduction is the transitive closure of $>$.

PROOF. Let aR_0b if a is a redex with contractum b, and let $>^*$ be the transitive closure of $>$. Reduction is the smallest transitive invariant relation containing R_0. By 1.4.(ii) $>^*$ is invariant. By 1.4(i) $>$ is reflexive and hence is easily seen to extend R_0. Hence $>^*$ extends reduction. But clearly reduction extends $>$ and hence $>^*$. So reduction is $>^*$.

1.6 The following formal system for $>$ will be very useful. The axioms are $a>a$ for any expression a. The axiom $a > a$ is strict if a is a variable. The rules are

$$\frac{a_1 > b_1 \ \ldots \ a_m > b_m}{a > b}$$

whenever $F(a_1,\ldots, a_m) = a$ and $F(b_1,\ldots, b_m) \doteq b$. Here $b' \doteq b$ if $b' = b$ or b' is a redex with contractum b.

It should be clear that $a > b$ iff $a > b$ has a proof in the above formal system. Moreover the proof can be chosen so that all its axioms are strict.

1.7 With the next definition we will be able to state our main result. A contraction operation is coherent if, for every redex $F(a_1,\ldots, a_m)$ with contractum a, $a_1 > b_1,\ldots, a_m > b_m$ implies that $F(b_1,\ldots, b_m)$ is a redex with contractum b such that $a > b$.

THEOREM

If the contraction operation is coherent then reduction has CRP.

PROOF. By lemmas 1.4.(iii) and 1.5 it suffices to show that $>$ has CRP. By induction on the proofs of $a > b$, $a > c$ we find an expression d such that $b > d$ and $c > d$.

If $a > b$ is an axiom then $a = b$ and we can let d be c. Similarly if $a > c$ is an axiom we can let d be b. Hence we can assume that $a > b$ and $a > c$ have proofs ending with inferences

$$\frac{a_1 > b_1 \ \ldots \ a_m > b_m}{a > b} \quad \text{and} \quad \frac{a_1 > c_1 \ \ldots \ a_m > c_m}{a > c}$$

respectively, where $F(a_1,\ldots, a_m) = a$, $F(b_1,\ldots, b_m) \doteq b$ and $F(c_1,\ldots, c_m) \doteq c$. (Note that because of our abuse of notation a proof of $a > c$ could end with an inference having premisses $a'_1 > c'_1 \ldots a'_m > c'_m$ where $F(a'_1,\ldots, a'_m) = a$ and $F(c'_1,\ldots, c'_m) \doteq c$. But by relabelling the variables in the proof it can be put in the required form.)

By induction hypothesis there are $d_1,\ldots, d_m$ such that $b_i > d_i$ and $c_i > d_i$ for $i = 1,\ldots, m$. Now define d as follows. If $F(d_1,\ldots, d_m)$ is a redex let d be its contractum, otherwise let d be $F(d_1,\ldots, d_m)$ itself. In either case $F(d_1,\ldots, d_n) \doteq d$, so that $F(b_1,\ldots, b_n) > d$ and $F(c_1,\ldots, c_n) > d$.

If $F(b_1,\ldots, b_n)$ is a redex then its contractum is b and as the contraction operation is coherent $b > d$. But if $F(b_1,\ldots, b_n)$ is not a redex then it is equal to b so that $b > d$ again. So $b > d$ and similarly $c > d$ as required.

1.8 The lemma below will be useful when we come to applying our main theorem. If $\vec{x} = x_1,\ldots, x_k$ is a list of variables and $\vec{c} = c_1,\ldots, c_k$ is a list of expressions then for any expression a $a[\vec{c}/\vec{x}]$ denotes the result of simultaneously substituting c_i for x_i in a for $i = 1,\ldots, k$ (making changes in the bound variables, where necessary, to avoid clashes). A contraction operation is <u>substitution preserving</u> if for all $\vec{c}, \vec{x}$ as above if a is a redex with contractum b then $a[\vec{c}/\vec{x}]$ is a redex with contractum $b[\vec{c}/\vec{x}]$. A binary relation R on expressions is <u>substitution preserving</u> if for all $\vec{c}, \vec{x}$ as above and $\vec{d} = d_1,\ldots, d_k$ aRb and $c_1Rd_1,\ldots, c_kRd_k$ implies that $a[\vec{c}/\vec{x}]Rb[\vec{d}/\vec{x}]$.

<u>LEMMA</u>

If the contraction operation is substitution preserving then so is $>$.

<u>PROOF</u>. Let $c_i > d_i$ for $i = 1,\ldots, k$. If $a > b$ then it has a proof and we can assume that all the axioms are strict. Replace each node $a' > b'$ of the proof by $a'[\vec{c}/\vec{x}] > b'[\vec{c}/\vec{x}]$, where if $a' > b'$ is an axiom $x_i > x_i$ it should be replaced by a proof of $c_i > d_i$. The result is easily seen to be a proof of $a[\vec{c}/\vec{x}] > b[\vec{d}/\vec{x}]$, using the assumption that the contraction operation is substitution preserving to show that a rule in the original proof

remains a rule under the above transformation.

§2. The lambda calculus

2.1 We show how to apply our main theorem to the lambda calculus. This is the expression system having the two forms APP and λ with arities 0, 0 and 1 respectively. We shall follow the usual convention and write $a(b)$ and $(\lambda x)a$ instead of $APP(a, b)$ and $\lambda((x)a)$ respectively. The β-contraction operation has redexes $((\lambda x)a)(b)$ with contracti $a[b/x]$.

THEOREM

β-contraction is coherent and hence β-reduction has CRP.

PROOF. First observe that β-contraction is substitution preserving. For if e is a redex $((\lambda x)a)(b)$ then $e[\vec{c}/\vec{x}]$ is $((\lambda x)a[\vec{c}/\vec{x}])(b[\vec{c}/\vec{x}])$ (as long as x is chosen so that it does not appear in $\vec{x}$ and is not free in b, c), and the latter is a redex with contractum $a[\vec{c}/\vec{x}]\left[b[\vec{c}/\vec{x}]/x\right]$, which is the same as $a[b/x][\vec{c}/\vec{x}]$. So by lemma 1.8 $>$ is substitution preserving.

To show coherence, let $((\lambda x)a)(b)$ be a redex and let $(\lambda x)a > a'$, $b > b_1$. We must show that $a'(b_1)$ is a redex, i.e. a' has the form $(\lambda x)a_1$, and then show that $a[b/x] > a_1[b_1/x]$.

Now consider a proof of $(\lambda x)a > a'$. We may assume that all the axioms are strict so that the proof must end with a rule $\dfrac{a > a_1}{(\lambda x)a > a'}$ where $(\lambda x)a_1 \doteq a'$. As $(\lambda x)a_1$ is not a redex $a' = (\lambda x)a_1$. As $>$ is substitution preserving and $b > b_1$, $a > a_1$ it follows that $a[b/x] > a_1[b_1/x]$ as required.

2.2 η-contraction is obtained from β-contraction by adding redexes $(\lambda x)(a(x))$ (with x not free in a) having contracti a. η-reduction has the CRP but the proof of this requires a slight modification of the argument of §1. Call the redexes for β-contraction β-redexes and the new redexes η-redexes. Let $>'$ be the smallest invariant relation such that (i) $a_1 >' b_1$, $a_2 >' b_2$ implies $a_1[a_2/x] >' b_1[b_2/x]$ and (ii) $a >' b$ implies $(\lambda x)(a(x)) >' b$, if x is not free in a. As in the proof of lemma 1.5 it is not hard to see that η-reduction is the transitive closure of $>'$. As in the proof of theorem 2.1 $>'$ can be shown to be substitution preserving, and hence β-contraction is coherent relative to $>'$, i.e. if $((\lambda x)a)(b)$ is a β-redex and $(\lambda x)a >' a'$, $b >' b_1$ then $a'(b_1)$ is a redex with a contractum c such that $a[b/x] >' c$.

Note that a formal system for $>'$ can be set up using the axioms and rules of the formal system for $>$ obtained from β-contraction and the additional η-rules $\dfrac{a >' b}{(\lambda x)(a(x)) >' b}$ if x is not free in a.

We are now ready to prove

THEOREM

$>'$ has CRP and hence so does η-reduction.

PROOF. The proof follows the pattern of the proof of theorem 1.7. By induction on proofs of $a >' b$, $a >' c$ we find d such that $b >' d$, $c >' d$. If neither proofs of $a >' b$, $a >' c$ end with an η-rule then we may argue as before. The new case to consider is when at least one of the proofs does end with an η-rule. By symmetry we may assume that the proof of $a >' b$ ends with an η-rule $\dfrac{a_1 >' b}{a >' b}$ where a is $(\lambda x)(a_1(x))$ and x is not free in a_1. There are two subcases to consider. First suppose that the proof of $a >' c$ also ends with an η-rule $\dfrac{a_1 >' c}{a >' b}$. In this case we can use the induction hypothesis to find d such that $b >' d$, $c >' d$ and we are done. It remains to consider the subcase where the proof of $a >' c$ ends with a rule $\dfrac{a_1(x) >' c_1}{a >' c}$ where $(\lambda x)c_1 \doteq c$. As $(\lambda x)c_1$ is not a β-redex $c = (\lambda x)c_1$. If $a_1(x) >' c_1$ is an axiom then $c_1 = a_1(x)$ and hence $a >' c$ is an axiom and we are under a previously considered case. Otherwise $a_1(x) >' c_1$ must have a proof ending with a rule $\dfrac{a_1 >' c_2 \quad x >' x}{a_1(x) >' c_1}$ where $c_2(x) \doteq c_1$. By induction hypothesis there is d such that $b >' d$ and $c_2 >' d$ and it only remains to show that $c >' d$. As $c_2(x) \doteq c_1$ either $c_2(x)$ is a β-redex with contractum c_1 or else $c_1 = c_2(x)$. Using the fact that a_1 does not contain x free and that $>'$ is substitution preserving, we can assume that c_2 does not contain x free. So in the first case c_2 is $(\lambda x)c_1$, so that $c = (\lambda x)c_1 = c_2$ and hence $c >' d$. In the second case $c = (\lambda x)c_1 = (\lambda x)(c_2(x))$ is an η-redex and as $c_2 >' d$ we get $c >' d$.

§3. Consistent sets of contraction schemes

3.1 We wish to generalise the example of the lambda calculus to a wide class of contraction operations defined in terms of schemes. To formulate these schemes we need to introduce k-ary metavariables for simple forms having arity $\overbrace{0, 0, \ldots, 0}^{k}$ for $k \geqslant 0$. Using these metavariables metaexpressions are generated using the same rules as for generating expressions, treating k-ary metavariables just like simple forms of arity $\overbrace{0, \ldots, 0}^{k}$.

Let H be a metaexpression and suppose $(\vec{x}_i)a_i$ is assigned to each k_i-ary metavariable Z_i occuring in H, where a_i is an expression and $\vec{x}_i$ is a list of k_i variables. Then $H[(\vec{x}_i)a_i/Z_i]_i$ denotes the result of repeatedly by replacing each subexpression $Z_i(H_1, \ldots, H_{k_i})$ of H by $a_i[H_1, \ldots, H_{k_i}/\vec{x}_i]$ for each Z_i. Note that the result will be an expression. We shall need the following lemma.

LEMMA

Let R be an invariant, substitution preserving relation. Let H be a metaexpression and let $(\vec{x}_i)a_i$ and $(\vec{x}_i)b_i$ be assigned to each metavariable Z_i occurring in H. Then $a_i R b_i$ for each Z_i implies $\bar{H} R \bar{\bar{H}}$, where $\bar{H} = H[(\vec{x}_i)a_i/Z_i]_i$ and $\bar{\bar{H}} = H[(\vec{x}_i)b_i/Z_i]_i$.

PROOF. Assume that $a_i R b_i$ for each Z_i. We prove that $\bar{H} R \bar{\bar{H}}$ by induction on the way that H is built up.

If H is a variable x then $\bar{H} R \bar{\bar{H}}$ is just xRx which is true by the invariance of R.

If H is $F(H_1, \ldots, H_m)$ then $\bar{H}$ is $F(\bar{H}_1, \ldots, \bar{H}_m)$ and $\bar{\bar{H}}$ is $F(\bar{\bar{H}}_1, \ldots, \bar{\bar{H}}_m)$. By induction hypothesis $\bar{H}_j R \bar{\bar{H}}_j$ for $j = 1, \ldots, m$ and hence, by the invariance of R, $\bar{H} R \bar{\bar{H}}$.

If H is $Z_i(H_1, \ldots, H_{k_i})$ then $\bar{H}$ is $a_i[\bar{H}_1, \ldots, \bar{H}_{k_i}/\vec{x}_i]$ and $\bar{\bar{H}}$ is $b_i[\bar{\bar{H}}_1, \ldots, \bar{\bar{H}}_{k_i}/\vec{x}_i]$. By induction hypothesis $\bar{H}_j R \bar{\bar{H}}_j$ for $j = 1, \ldots, k_i$. Also $a_i R b_i$, so that as R is substitution preserving $\bar{H} R \bar{\bar{H}}$.

3.2 A contraction scheme is a pair (H, H') of metaexpressions having no free variables such that:

(i) The metavariables occurring in H, H' are Z_{ij} for

$i = 1, \ldots, m$, $j = 1, \ldots, m_i$

(ii) H is an expression $F((\vec{x}_1)H_1, \ldots, (\vec{x}_m)H_m)$ where for each i either

(a) H_i is $Z_{i1}(\vec{x}_i)$ and $m_i = 1$, or

(b) H_i is $G_i((\vec{y}_{i1})Z_{i1}(\vec{x}_i\vec{y}_{i1}), \ldots, (\vec{y}_{im_i})Z_{im_i}(\vec{x}_i\vec{y}_{im_i}))$

F is called the defined form of the scheme and each G_i is a primitive form of the scheme.

An instance of the scheme (H, H') is a pair (a, b) such that for some assignment $(\vec{x}_i\vec{y}_{ij})a_{ij}$ to each metavariable Z_{ij} $a = H[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$ and $b = H'[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$. If (a) holds for a given i then $j = 1$ and $\vec{y}_{ij}$ denotes the empty sequence. Note that if (a, b) is an instance of (H, H') then given a choice of the $\vec{x}_i\vec{y}_{ij}$ the a_{ij} can be determined from H and a only.

A set of contraction schemes $\mathcal{C}$ is consistent if

(i) no defined form of any $\mathcal{C}$-scheme is also a primitive form of a $\mathcal{C}$-scheme.

(ii) if two distinct $\mathcal{C}$-schemes have the same defined form then there is an i such that in each of the schemes a primitive form G_i is assigned but these forms are different in the two schemes.

If (a, b), (a, b') are instances of $\mathcal{C}$-schemes where $\mathcal{C}$ is a consistent set of contraction schemes it follows from (ii) that they must both be instances of the same $\mathcal{C}$-scheme (H, H'). Moreover given a choice of the $\vec{x}_i\vec{y}_{ij}$, the a_{ij} so that $a = H[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$ are uniquely determined and hence both b and b' must be $H'[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$. So $b = b'$. It follows from this that the set of instances (a, b) of $\mathcal{C}$-schemes, where $\mathcal{C}$ is a consistent set of contraction schemes, form the graph of a contraction operator.

THEOREM

Given a consistent set $\mathcal{C}$ of contraction schemes the resulting contraction operation is coherent.

PROOF. Let $F(a_1, \ldots, a_n)$ be a redex with contractum a and let $a_1 > b_1$,

..., $a_m > b_m$. We must show that $F(b_1, ..., b_m)$ is a redex and has contractum b where $a > b$.

By assumption $(F(a_1, ..., a_m), a)$ is an instance of a contraction scheme (H, H') from the given set $\mathcal{C}$. H is $F((\vec{x}_1)H_1, ..., (\vec{x}_m)H_m)$ and $F(a_1, ..., a_m)$ is $H[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$ and a_i is $H_i[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_j$. So a_i is a_{i1} if H_i is $Z_{i1}(\vec{x}_i)$ and $G_i(a_{i1}, ..., a_{im_i})$ otherwise. In the first case let $b_{i1} \doteq b_i$. In the second case, as $a_i > b_i$, it has a proof and, as we can assume that all axioms are strict, it must end with an inference

$$\frac{a_{i1} > b_{i1} \quad \dots \quad a_{im_i} > b_{im_i}}{a_i > b_i}$$

where $G_i(b_{i1}, ..., b_{im_i}) \doteq b_i$.

As G_i is a primitive form and hence cannot be a defined form it follows that $G_i(b_{i1}, ..., b_{im_i})$ cannot be a redex and hence it is the same as b_i. We have now defined b_{ij} for each ij so that $a_{ij} > b_{ij}$ and $b_i = H_i[(\vec{x}_i\vec{y}_{ij})b_{ij}/Z_{ij}]_j$. Now if we let $b = H'[(\vec{x}_i\vec{y}_{ij})b_{ij}/Z_{ij}]_{ij}$ then $(F(b_1, ..., b_m), b)$ is also an instance of the scheme (H, H') so that $F(b_1, ..., b_n)$ is a redex with contractum b. As $a = H'[(\vec{x}_i\vec{y}_{ij})b_{ij}/Z_{ij}]_{ij}$ and $a_{ij} > b_{ij}$ for all i, we may use lemma 3.1 to deduce that $a > b$ once we know that $>$ is substitution preserving. By lemma 1.8 it suffices to prove that contraction preserves substitution. So given $\vec{c}, \vec{x}$, let a be a redex with contractum b. Then (a, b) is an instance of some contraction scheme (H, H'). So $a = H[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$ and $b = H'[(\vec{x}_i\vec{y}_{ij})a_{ij}/Z_{ij}]_{ij}$. Then $a[\vec{c}/\vec{x}] = H[(\vec{x}_i\vec{y}_{ij})a_{ij}[\vec{c}/\vec{x}]/Z_{ij}]_{ij}$ and $b[\vec{c}/\vec{x}] = H'[(\vec{x}_i\vec{y}_{ij})a_{ij}[\vec{c}/\vec{x}]/Z_{ij}]_{ij}$ where we use the fact here that H and H' contain no free variables. Thus $(a[\vec{c}/\vec{x}], b[\vec{c}/\vec{x}])$ is also an instance of (H, H') and hence $a[\vec{c}/\vec{x}]$ is a redex with contractum $b[\vec{c}/\vec{x}]$.

3.3. The β-contraction for the lambda calculus, considered in §2, is obtained as above from the contraction scheme

$$(APP(\lambda((x)Z_{11}(x), Z_{21}), Z_{11}(Z_{21})).$$

We consider some further schemes.

Pairing. Let PAIR be a form of arity 0, 0 and let p, q be forms of arity 0. The pairing contraction has redexes p(PAIR(a, b)) or q(PAIR(a, b)) with contracta a or b respectively. This contraction comes from the consistent set of two contraction schemes

$$(p(PAIR(Z_{11}, Z_{12})), Z_{11})$$

and

$$(q(PAIR(Z_{11}, Z_{12})), Z_{12}).$$

Definition by cases. Let R_n be a form of arity $\overbrace{0,\ldots,0}^{n+1}$ $(n > 0)$ and let $1_n,\ldots, n_n$ be constants. The n-ary definition by cases contraction has redexes $R_n(1_n, a_1,\ldots, a_n)$ or ... or $R_n(n_n, a_1,\ldots, a_n)$ with contracta a_1 or ... or a_n respectively. This contraction comes from the consistent set of n contraction schemes:

$$(R_n(1_n, Z_{21},\ldots, Z_{n+11}), Z_{21})$$

...

$$(R_n(n_n, Z_{21},\ldots, Z_{n+11}), Z_{n+11}).$$

Primitive Recursion. Let 0 be a constant, s a form of arity 0 and R a form of arity 0, 0, 2. The primitive recursion contraction has redexes R(0, b, (x, y)c) or R(s(a), b, (x, y)c) with contracta b or c[a, R(a, b, (x, y)c)/x, y] respectively. This contraction comes from the consistent set of two contraction schemes:

$$(R(0, Z_{21}, (x, y)Z_{31}(x, y)), Z_{21}) \text{ and}$$

$$(R(s(Z_{11}), Z_{21}, (x, y)Z_{31}(x, y)), Z_{31}(Z_{11}, R(Z_{11}, Z_{21}, (x, y)Z_{31}(x, y))))$$

All the schemes considered above may be combined to form a set of schemes which is still clearly consistent. So our results show that the contraction obtained by using all the schemes has a reduction with CRP. As in §2 this still remains true when η-redexes and their contracta are included.

Manchester
July 1978.