Appendix A

Proof of the Church-Rosser Theorem

In this appendix, we will give a proof of the Church-Rosser Theorem. The proof we present is due to Tait and Martin-Löf, and is outlined in Barendregt[4] and in Amadio and Curien[3].

Recall the statement of the Church-Rosser Theorem:

Theorem A.1 (Church-Rosser) *Let E, E_1, and E_2 be λ-terms such that $E \to_\beta^* E_1$ and $E \to_\beta^* E_2$. Then there exists a λ-term E_3 such that, up to α-equivalence, $E_1 \to_\beta^* E_3$ and $E_2 \to_\beta^* E_3$.*

Throughout this discussion, we shall consider α-equivalent expressions to be equal. In particular, this means that we may assume, when proving a result, that any α-conversion necessary to prevent name clashes has already been done. This convention is standard in proofs in the λ-calculus, and it allows us to avoid the messy details associated with renaming variables in the middle of a reduction sequence. Thus, all of the results in this appendix are qualified with "up to α-equivalence." However, for the sake of clarity, we will point out some of the places where our convention is used as we go along.

The property of β-reduction (together with any necessary α-conversion) asserted by the Church-Rosser theorem is known as the *diamond property* (also *confluence*[1]):

Definition A.1 *A binary relation $\to$ has the* diamond property *if, whenever $a \to b$ and $a \to c$, there exists some d such that $b \to d$ and $c \to d$.*

The diamond property is illustrated in Figure A.1. The solid arrows indicate the assumed reductions $a \to b$ and $a \to c$. The dashed arrows indicate the reductions $b \to d$ and $c \to d$ that follow from these assumptions and the diamond property. A proof of the Church-Rosser Theorem is really a proof that $\to_\beta^*$ has the diamond property.

Lemma A.1 *If a binary relation $\to$ has the diamond property, then so do its transitive closure $\to^+$ and its reflexive, transitive closure $\to^*$.*

[1]Confluence and the diamond property are not quite equivalent notions. Strictly speaking, the statement that a relation $\to$ is confluent is equivalent to the statement that its reflexive, transitive closure, $\to^*$, satisfies the diamond property.

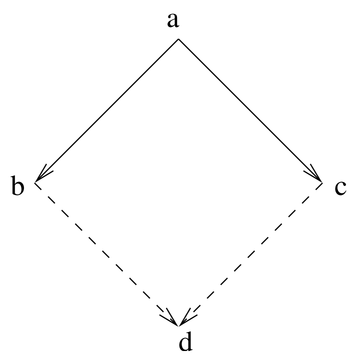
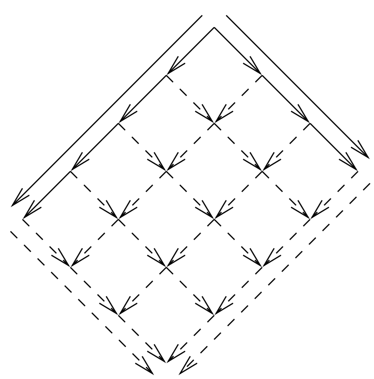


Figure A.1: The diamond property.

Proof The following diagram gives the intuition:



The short arrows represent reductions in $\rightarrow$. Solid arrows indicate reductions whose existence is part of the hypothesis of the diamond property (i.e., "*if* $a \rightarrow b$ and $a \rightarrow c \ldots$"), and dashed arrows indicate reductions whose existence is asserted by the diamond property (i.e., "*then* there exist reductions $b \rightarrow d$ and $c \rightarrow d$"). The long arrows indicate reductions in the transitive closure, $\rightarrow^+$. Essentially, we can view a reduction in $\rightarrow^+$ as a sequence of reductions in $\rightarrow$. By applying the diamond property repeatedly to the reductions in $\rightarrow$, we get reductions in $\rightarrow^+$ that complete the diamond in the transitive closure. For the reflexive, transitive closure, we have the additional case that a reduction in $\rightarrow^*$ may consist of no updates at all. But then if $a \rightarrow^* b$ and $a \rightarrow^* c$ with $a = b$, we can take $d = c$.$\square$

Therefore, if we can prove that $\rightarrow_\beta$ has the diamond property, then by Lemma A.1, we can conclude that $\rightarrow_\beta^*$ also has the diamond property. However, it turns out that $\rightarrow_\beta$ does not have the diamond property (exercise). So we will focus our efforts on finding another binary relation that does have the diamond property, and whose transitive closure is $\rightarrow_\beta^*$.

Definition A.2 *Define a binary relation $\twoheadrightarrow$ on λ-terms as follows:*

$$\frac{}{M \twoheadrightarrow M}\ (1) \qquad \frac{M \twoheadrightarrow M'}{\lambda x.M \twoheadrightarrow \lambda x.M'}\ (2)$$

$$\frac{M \twoheadrightarrow M' \quad N \twoheadrightarrow N'}{MN \twoheadrightarrow M'N'}\ (3) \qquad \frac{M \twoheadrightarrow M' \quad N \twoheadrightarrow N'}{(\lambda x.M)N \twoheadrightarrow M'[N'/x]}\ (4)$$

The relation $\twoheadrightarrow$ is a kind of "parallel reduction," under which certain redices may be reduced simultaneously. For example, this definition of reduction allows us to reduce both the rator and the rand of an application in a single step.

Before we consider the properties of $\twoheadrightarrow$, we will establish an important property of substitution, known as the Substitution Lemma:

Lemma A.2 (Substitution Lemma) *Let M, N, and P be λ-terms. If $x \neq y$ and $x \notin FV[P]$, then*

$$M[N/x][P/y] \quad = \quad M[P/y][N[P/y]/x]$$

Proof The intuition behind this lemma should be clear. On the left-hand side, the substitution $[P/y]$ replaces all occurrences of y in M and N with P. On the right-hand side, the substitutions are reversed. But the naive reversal, $M[P/y][N/x]$, would fail to replace any y's in N with P, and so we need to perform this additional substitution explicitly and write $M[P/y][N[P/y]/x]$. Formally, we will prove the result by induction on the structure of M. There are three cases:

1. M is a variable: if $M = x$, then the left-hand side is $x[N/x][P/y]$, which is just $N[P/y]$, and the right-hand side is $x[P/y][N[P/y]/x]$, which reduces to $x[N[P/y]/x]$, and then to $N[P/y]$. If $M = y$, then the left-hand side is P and the right-hand side is $P[N[P/y]/x]$. Since $x \notin FV[P]$, there are no x's to substitute in P, and so the right-hand side reduces to P. If M is some variable z, distinct from x and y, then both sides reduce to z. In any case the left-hand and right-hand sides are equal.

2. $M = \lambda z.M'$: we assume that the necessary α-conversion has been performed so that z is equal to neither x nor y, and furthermore that $z \notin FV[P]$. We then have

$$\begin{aligned} (\lambda z.M')[N/x][P/y] &= \lambda z.M'[N/x][P/y] \\ &= \lambda z.M'[P/y][N[P/y]/x] \text{ (by induction)} \\ &= (\lambda z.M')[P/y][N[P/y]/x] \end{aligned}$$

3. $M = M_1M_2$: then we have

$$\begin{aligned} (M_1M_2)[N/x][P/y] &= M_1[N/x][P/y]M_2[N/x][P/y] \\ &= M_1[P/y][N[P/y]/x]M_2[P/y][N[P/y]/x] \text{ (by induction)} \\ &= (M_1M_2)[P/y][N[P/y]/x] \end{aligned}$$

This completes the proof.□

Lemma A.3 *Let M and N be λ-terms with $N \twoheadrightarrow N'$. Then $M[N/x] \twoheadrightarrow M[N'/x]$.*

Proof We prove the result by induction on the structure of M. There are five cases to consider:

1. $M = x$: then we have $M[N/x] = x[N/x] = N \twoheadrightarrow N' = x[N'/x] = M'[N/x]$, and the result holds.
2. M is a variable other than x, say y: then $M[N/x] = y[N/x] = y \twoheadrightarrow y = y[N'/x] = M[N'/x]$, and the result holds.
3. $M = PQ$: then $M[N/x] = P[N/x]Q[N/x]$. By induction, $P[N/x] \twoheadrightarrow P[N'/x]$ and $Q[N/x] \twoheadrightarrow Q[N'/x]$. Therefore, by Rule (3) in Definition A.2, $P[N/x]Q[N/x] \twoheadrightarrow P[N'/x]Q[N'/x] = (PQ)[N'/x]$. Thus, $M[N/x] \twoheadrightarrow M[N'/x]$, and the result holds.
4. $M = \lambda x.P$: then $M[N/x] = M[N'/x] = M$, and the result holds trivially (recall that $(\lambda x.P)[N/x] = \lambda x.P$ for all P, N—we are attempting to substitute on the binding variable).
5. $M = \lambda y.P$, where y is some variable other than x: we assume that any necessary α-conversion has already been done, and we have $M[N/x] = \lambda y.P[N/x]$. By induction, $P[N/x] \twoheadrightarrow P[N'/x]$, and so by Rule (2) in Definition A.2, $M[N/x] = \lambda y.P[N/x] \twoheadrightarrow \lambda y.P[N'/x] = M[N'/x]$, and the result holds.

Note that case 4 is not strictly necessary, as we can absorb it into case 5 by assuming that the necessary α-conversion has been done.□

Lemma A.4 *Let $M \twoheadrightarrow M'$ and $N \twoheadrightarrow N'$. Then $M[N/x] \twoheadrightarrow M'[N'/x]$.*

Proof The proof is by induction on the structure of the statement $M \twoheadrightarrow M'$. By examining the Post rules in Definition A.2, we see that this statement could take one of four forms:

1. $M = M'$ (i.e. $M \twoheadrightarrow M'$ comes from Rule (1)): by Lemma A.3, $M[N/x] \twoheadrightarrow M[N'/x] = M'[N'/x]$.
2. $M = \lambda y.M_1$, $M' = \lambda y.M_1'$, with $M_1 \twoheadrightarrow M_1'$ (Rule (2)): we assume that any necessary α-conversion has already been done. Then $(\lambda y.M_1)[N/x] = \lambda y.M_1[N/x]$. By induction, $M_1[N/x] \twoheadrightarrow M_1'[N'/x]$, and we have $M[N/x] = \lambda y.M_1[N/x] \twoheadrightarrow \lambda y.M_1'[N'/x] = M'[N'/x]$, and the result holds.
3. $M = PQ$, $M' = P'Q'$, with $P \twoheadrightarrow P'$ and $Q \twoheadrightarrow Q'$ (Rule (3)): then $M[N/x] = P[N/x]Q[N/x]$. By induction, $P[N/x] \twoheadrightarrow P'[N'/x]$ and $Q[N/x] \twoheadrightarrow Q'[N'/x]$; we then have

$$\begin{aligned} M[N/x] &= P[N/x]Q[N/x] \\ &\twoheadrightarrow P'[N'/x]Q'[N'/x] \text{ (by Rule (3))} \\ &= M'[N'/x], \end{aligned}$$

   and the result holds.

4. $M = (\lambda y.P)Q$, $M' = P'[Q'/y]$, with $P \twoheadrightarrow P'$ and $Q \twoheadrightarrow Q'$ (Rule (4)): then $M[N/x] = (\lambda y.P[N/x])Q[N/x]$. By induction, $P[N/x] \twoheadrightarrow P'[N'/x]$ and $Q[N/x] \twoheadrightarrow Q'[N'/x]$. Hence $\lambda y.P[N/x] \twoheadrightarrow \lambda y.P'[N'/x]$; we then have

$$\begin{array}{rcl} M[N/x] & = & (\lambda y.P[N/x])Q[N/x] \\ & \twoheadrightarrow & P'[N'/x][Q'[N'/x]/y] \text{ (by Rule (4))} \\ & = & P'[Q'/y][N'/x] \text{ (by the Substitution Lemma)} \\ & = & M'[N'/x], \end{array}$$

   and the result holds.

This completes the proof.□

Lemma A.5 *If $\lambda x.M \twoheadrightarrow N$, then $N = \lambda x.M'$ with $M \twoheadrightarrow M'$.*

Proof We prove the result by case analysis on the structure of the statement $\lambda x.M \twoheadrightarrow N$. By pattern-matching on the Post rules in Definition A.2, we see that two cases are possible:

1. $\lambda x.M = N$ (i.e. $\lambda x.M \twoheadrightarrow N$ comes from Rule (1)): then the result holds trivially, as $M \twoheadrightarrow M$ and $N = \lambda x.M$.

2. $N = \lambda x.M'$ with $M \twoheadrightarrow M'$ (i.e. $\lambda x.M \twoheadrightarrow N$ comes from Rule (2)): this is exactly what we set out to prove, so we are done.

This completes the proof.□

Lemma A.6 *If $MN \twoheadrightarrow L$, then either $L = M'N'$ with $M \twoheadrightarrow M'$, $N \twoheadrightarrow N'$ or $M = \lambda x.P$, $L = P'[N'/x]$, $P \twoheadrightarrow P'$, $N \twoheadrightarrow N'$.*

Proof We prove the result by case analysis on the structure of the statement $MN \twoheadrightarrow L$. By pattern-matching on the Post rules in Definition A.2, we see that three cases are possible:

1. $MN = L$ (i.e. $MN \twoheadrightarrow L$ comes from Rule (1)): then the result holds trivially, since $M \twoheadrightarrow M$, $N \twoheadrightarrow N$, and $L = MN$.

2. $L = M'N'$, $M \twoheadrightarrow M'$, and $N \twoheadrightarrow N'$ (i.e. $MN \twoheadrightarrow L$ comes from Rule (3)): this is exactly what we wish to prove, so we are done.

3. $M = \lambda x.M_1$, $L = M_1'[N'/x]$, $M_1 \twoheadrightarrow M_1'$, and $N \twoheadrightarrow N'$ (i.e. $MN \twoheadrightarrow L$ comes from Rule (4)): again, this is exactly what we set out to prove (take $P = M_1$).

This completes the proof.□

Lemma A.7 *$\twoheadrightarrow$ has the diamond property.*

Proof Let M, M_1, and M_2 be λ-terms such that $M \twoheadrightarrow M_1$ and $M \twoheadrightarrow M_2$. We claim that there exists a λ-term M_3 such that $M_1 \twoheadrightarrow M_3$ and $M_2 \twoheadrightarrow M_3$. We prove our claim by induction on the structure of the expression $M \twoheadrightarrow M_1$. By the Post rules in Definition A.2, $M \twoheadrightarrow M_1$ could take one of four forms:

1. $M = M_1$ (i.e. $M \twoheadrightarrow M_1$ comes from Rule (1)): then we can take $M_3 = M_2$ and we are done (we have $M_2 \twoheadrightarrow M_3$ because $M_2 = M_3$, and also $M_1 = M \twoheadrightarrow M_2 = M_3$ gives $M_1 \twoheadrightarrow M_3$).

2. $M = PQ$, $M_1 = P'Q'$, $P \twoheadrightarrow P'$, and $Q \twoheadrightarrow Q'$ (Rule (3)): this is the case where M is an application, and M_1 is the same application, in which the rator and/or rand has been reduced, but the rator has not been applied to the rand. Here, by Lemma A.6 applied to $M \twoheadrightarrow M_2$, there are two subcases:

   (a) $M_2 = P''Q''$ with $P \twoheadrightarrow P''$ and $Q \twoheadrightarrow Q''$: we then have $P \twoheadrightarrow P'$ and $P \twoheadrightarrow P''$. Furthermore, $Q \twoheadrightarrow Q'$ and $Q \twoheadrightarrow Q''$. So by induction, there exist λ-terms P''' and Q''' such that $P' \twoheadrightarrow P'''$, $P'' \twoheadrightarrow P'''$, $Q' \twoheadrightarrow Q'''$, and $Q'' \twoheadrightarrow Q'''$. Now, we have $M_1 = P'Q' \twoheadrightarrow P'''Q'''$ (Rule (3)) and $M_2 = P''Q'' \twoheadrightarrow P'''Q'''$ (Rule (3)), so we may take $M_3 = P'''Q'''$, and we are done.

   (b) $P = \lambda x.P_1$, $M_2 = P_1''[Q''/x]$, $P_1 \twoheadrightarrow P_1''$, and $Q \twoheadrightarrow Q''$: then, by Lemma A.5, we have $P' = \lambda x.P_1'$ with $P_1 \twoheadrightarrow P_1'$. We now have $P_1 \twoheadrightarrow P_1'$ and $P_1 \twoheadrightarrow P_1''$. Further, $Q \twoheadrightarrow Q'$ and $Q \twoheadrightarrow Q''$. Now, by induction, there exist λ-terms P_1''' and Q''' such that $P_1' \twoheadrightarrow P_1'''$, $P_1'' \twoheadrightarrow P_1'''$, $Q' \twoheadrightarrow Q'''$, and $Q'' \twoheadrightarrow Q'''$. Finally, $M_1 = P'Q' = (\lambda x.P_1')Q' \twoheadrightarrow P_1'''[Q'''/x]$ (Rule (4)) and $M_2 = P_1''[Q''/x] \rightarrow P_1'''[Q_1'''/x]$ (Lemma A.4), and so we may take $M_3 = P_1'''[Q_1'''/x]$, and we are done.

3. $M = (\lambda x.P)Q$, $M_1 = P'[Q'/x]$, $P \twoheadrightarrow P'$, and $Q \twoheadrightarrow Q'$ (Rule (4)): this is the case where M is an application, and M' is the term resulting from substituting the rand into the rator (possibly after first reducing the rand or the body of the rator). By Lemma A.6 applied to $M \twoheadrightarrow M_2$, there are two subcases:

   (a) $M_2 = (\lambda x.P'')Q''$, $P \twoheadrightarrow P''$, and $Q \twoheadrightarrow Q''$: we have $P \twoheadrightarrow P'$ and $P \twoheadrightarrow P''$. Further, $Q \twoheadrightarrow Q'$ and $Q \twoheadrightarrow Q''$. Thus, by induction, there exist λ-terms P''' and Q''' such that $P' \twoheadrightarrow P'''$, $P'' \twoheadrightarrow P'''$, $Q' \twoheadrightarrow Q'''$, and $Q'' \rightarrow Q'''$. Now, we have $M_1 = P'[Q'/x] \twoheadrightarrow P'''[Q'''/x]$ (Lemma A.4), and $M_2 = (\lambda x.P'')Q'' \twoheadrightarrow P'''[Q'''/x]$ (Rule (4)). Hence, we may take $M_3 = P'''[Q'''/x]$.

   (b) $M_2 = P''[Q''/x]$, $P \twoheadrightarrow P''$, $Q \twoheadrightarrow Q''$: we have $P \twoheadrightarrow P'$ and $P \twoheadrightarrow P''$. Also, $Q \twoheadrightarrow Q'$ and $Q \twoheadrightarrow Q''$. By induction, there exist λ-terms P''' and Q''' with $P' \twoheadrightarrow P'''$, $P'' \twoheadrightarrow P'''$, $Q' \twoheadrightarrow Q'''$, and $Q'' \rightarrow Q'''$. Then $M_1 = P'[Q'/x] \twoheadrightarrow P'''[Q'''/x]$ (Lemma A.4) and $M_2 = P''[Q''/x] \twoheadrightarrow P'''[Q'''/x]$ (Lemma A.4 again). Thus, we can take $M_3 = P'''[Q'''/x]$.

4. $M = \lambda x.P$, $M_1 = \lambda x.P'$, $P \twoheadrightarrow P'$ (Rule (2)): by Lemma A.5, we have $M_2 = \lambda x.P''$ for some λ-term P'' with $P \twoheadrightarrow P''$. By induction, there exists a λ-term P''' with $P' \twoheadrightarrow P'''$ and $P'' \twoheadrightarrow P'''$. Then $M_1 = \lambda x.P' \twoheadrightarrow \lambda x.P'''$ and $M_2 = \lambda x.P'' \twoheadrightarrow \lambda x.P'''$ (by Rule (2)). Therefore, we may take $M_3 = \lambda x.P'''$.

This completes the proof.□

We are now ready to complete the proof of the Church-Rosser Theorem:

Proof of the Church-Rosser Theorem By Lemma A.7, $\twoheadrightarrow$ has the diamond property. Since $\rightarrow_\beta \subseteq \twoheadrightarrow \subseteq \rightarrow^*_\beta$ (exercise: show that these inclusions are proper; that is, show that $\rightarrow_\beta \subset \twoheadrightarrow \subset \rightarrow^*_\beta$), and $\rightarrow^*_\beta$ is the reflexive, transitive closure of $\rightarrow_\beta$, it follows that $\rightarrow^*_\beta$ is the reflexive, transitive closure of $\twoheadrightarrow$ (actually, it is also just the transitive closure of $\twoheadrightarrow$, since $\twoheadrightarrow$ is already reflexive). Then by Lemma A.1, $\rightarrow^*_\beta$ has the diamond property, and we are done.$\square$