LABORATORY FOR COMPUTER SCIENCE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

# Notes on the Confluence Property of Terms Rewriting Systems and the $\lambda$-calculus

Computation Structures Group Memo 321

**Zena M. Ariola**

## **** DRAFT ****

This report describes research done at the Laboratory for Computer Science of the Massachusetts Institute of Technology. Funding for this work has been provided in part by the Advanced Research Projects Agency of the Department of Defense under the Office of Naval Research contract N00014-89-J-1988 (MIT) and N0039-88-C-0163 (Harvard).

545 TECHNOLOGY SQUARE, CAMBRIDGE, MASSACHUSETTS 02139

# Notes on the Confluence Property of Terms Rewritings Systems and the $\lambda$-calculus

*Zena M. Ariola*
Aiken Computational Laboratory
Harvard University

November 28, 1990

In this document we will prove the Church-Rosser theorem for both Regular[1] Term Rewriting Systems (TRS's) and the $\lambda$-calculus. There are many ways to prove the Church-Rosser theorem.

## 1 Basic Definition and Properties

### 1.1 Order

**Definition 1.1** *Let $R$ be a binary relation on a set $X$. We say*
- *$R$ is reflexive if* $\forall\, x \in X\ [x\ R\ x]$
- *$R$ is antisymmetric if* $\forall\, x, y \in X\ [x\ R\ y\ \wedge\ y\ R\ x \implies x = y]$
- *$R$ is transitive if* $\forall\, x, y, z \in X\ [x\ R\ y\ \wedge\ y\ R\ z \implies x\ R\ z]$
- *$R$ is trichotomous if* $\forall\, x, y \in X\ [x\ R\ y\ \vee\ y\ R\ x\ \vee\ x = y]$

**Definition 1.2 (Partial Order)** *A partial order in a set $X$ is a reflexive, antisymmetric, and transitive relation in $X$.*

It is customary to use the symbol "$\leq$" for a partial order. The reason for the qualification "partial" is that some questions about order may be left unanswered.

**Definition 1.3 (Linear Order)** *A linear (total) order in a set $X$ is a trichotomous partial order in $X$.*

[1] Researchers have coined the word "*Orthogonal*" for this subclass of TRS's. However, in this document we will play conservative and still use the widely known term "Regular".

A linear order is frequently called a *chain.*

Example: The most natural example of a partial (and not total) order is the inclusion relation, $\subseteq$, on the power set, $P(X)$, of a set $X$. $\subseteq$ is a total order iff $X$ is either empty or is a singleton set. An example of a total order is the relation "less than or equal to" in the set of natural numbes.

**Definition 1.4 (Partially Ordered Set)** *A partially order set $\mathcal{X}$ is an ordered pair $(X, \leq)$, where $X$ is a set and $\leq$ is a partial order in $X$.*

In the following we will not make the distinction between a partially ordered set and the domain of the partial order, that is, we will use $\mathcal{X}$ and $X$ interchangeably.

**Definition 1.5 (Minimal element)** *A minimal element of a partially order set $(X, \leq)$ is a $y$ in $X$ such that*

$$\nexists\, x \in X\, [x \;\leq\; y]$$

A minimal element need not be unique (contrast wiiith the definition of a least element, given in the appendix).

Exercise: Does the set of all non-empty subsets of a non-empty set $X$ have either a minimum element or a least element or both?

**Definition 1.6 (Initial segment)** *Let $X$ be a partially order set, if $x \in X$, then*

$$\{y \in X \mid y \;<\; x\}$$

*is called the initial segment of $x$; we shall denote it by $s(x)$.*

Note that in the above definition we used the symbol $<$ and not $\leq$. For this reason the above is usually called the "strict" initial segment of $x$.

**Definition 1.7 (Minimum condition)** *A partially ordered set $X$ satisfies the minimum condition if each non-empty subset of $X$ has a minimal element.*

We now state the following theorem without proof.

**Theorem 1.8 (Noetherian induction)** *Given $S$, a subset of a partially ordered set $X$ with minimum condition,*

$$[\forall\, x \in X,\;\; s(x) \;\subseteq\; S \;\Longrightarrow\; x \in S] \;\;\Longrightarrow\;\; S = X$$

Notice that in the above we didn't make any assumption about a starting element. This is so because all the minimal elements of $X$ are included in $S$ by definition. Indeed, if $x$ is the minimun element of $X$, $s(x)$ is empty and therefore $s(x) \;\subseteq\; S$, then $x \in S$.

At this point the reader may feel confused about the difference among the various kind of "induction principles" he may have come across. In the following we will try to trow some light on these differences, if any. We will consider mathematical induction, structural induction and transfinite induction (defined in the appendix).
Let's first say that Noetherian induction is the general version of "structural induction". Structural induction, as the name may recall, consists in reasoning on the structure of a term or a formula. For example, most of the proofs in propositional logic, goes like this

$$\text{suppose } \phi \text{ and } \psi \text{ are true, then prove that } \phi \wedge \psi \text{ is true}$$

We clearly have a partial order set (*i.e.* $\phi \leq \phi \wedge \psi$ and $\psi \leq \phi \wedge \psi$), which satisfy the minimum condition, where the minimum elements are the atomic terms or formulae.

While, the main difference beween Noetherian induction and both mathematical induction and transfinite induction, is that the first is defined on partially order sets, while both mathematical induction and transfinite induction are defined on well-orders. This means that, noetherian iinduction only requires that each chain in $X$ has a least element, and not an arbitrary subset of $X$. Moreover, Noetherian induction, like transfinite induction, passes to each element from the set of its predeccesors, and, as said before, does not make any assumption about a starting element.

For a definition of both transfinite induction and least element the reader may refer to the appendix.

## 1.2 Application of Noetherian Induction

We can think of the reduction relation as extablishing an ordering between terms. For example, you can read "$x \twoheadrightarrow y$" as saying "$y \leq x$". It is left to the reader to check that the relation "$\leq$" is a partial order. We would like to have a tool that allows us to reason about terms. In order to do that we have to require that the reduction relation be SN, otherwise, we would not have the minimun element property.

**Definition 1.9 (Complete)** *Let P be a predicate defined on a partially order set $X$. We say that P is complete iff*

$$\forall\, x \in X,\ \ \forall\, y \in \{y \mid x \stackrel{+}{\twoheadrightarrow} y\},\ \ P(y) \implies P(x)$$

We introduce a property of a predicate which says that a predicate is complete if it holds for an arbitrary element $x$ of $X$ whenever P holds for all the elements less defined than $x$.

**Theorem 1.10** *Given TRS $(X,R)$, if R is noetherian and P is a complete predicate then*

$$\forall\, x \in X,\ \ P(x)$$

*Proof:* Suppose, by contradiction, that P does not hold in each element of $X$, therefore, the set, $S$, of all element which do not satisfy P is non-empty, and by the minimum condition

$$\exists\, m \in S,\ \ \forall\, z \in S,\ m \leq z$$

This means that $s(m) \not\subseteq S$ (otherwise $m$ would not be a minimum). We then have

$$s(m) \not\subseteq S \wedge m \in S$$

This contradicts the hypothesis that P is a complete predicate. ∎

**Lemma 1.11 (Newman's Lemma)** $SN \wedge WRC \Longrightarrow CR$ .

*Proof:* We want to prove that the predicate $P(x)$ defined below is complete

$$P(x): \ \forall\, y, z,\ [x \twoheadrightarrow y \ \wedge \ x \twoheadrightarrow z \ \Longrightarrow\ \exists\, s \text{ such that } y \twoheadrightarrow s \ \wedge \ z \twoheadrightarrow s]$$

Without loss of generality assume that

$$x \longrightarrow y_1 \twoheadrightarrow y \ \wedge \ x \longrightarrow z_1 \twoheadrightarrow z$$

By WCR:

$$\exists\, u,\ y_1 \twoheadrightarrow u \ \wedge \ z_1 \twoheadrightarrow u$$

By induction hypothesis P holds in $y_1$ (since $y_1 \leq x$),

$$\exists\, v,\ y \twoheadrightarrow v \ \wedge \ u \twoheadrightarrow v$$

By induction hypothesis P holds in $z_1$ (since $z_1 \leq x$),

$$\exists\, t,\ v \twoheadrightarrow t \ \wedge \ z \twoheadrightarrow t$$

Thus proving P($x$). See the diagram below:

$$\begin{tikzcd}
x \arrow[r] \arrow[d] \arrow[dr, phantom, "WCR"] & z_1 \arrow[r] \arrow[d, two heads] & z \arrow[dd, two heads] \\
y_1 \arrow[r, two heads] \arrow[d, two heads] \arrow[dr, phantom, "HP"] & u \arrow[d, two heads] \arrow[r, phantom, "HP"] & \\
y \arrow[r, two heads] & v \arrow[r, two heads] & t
\end{tikzcd}$$

∎

## 1.3 Reduction Properties

Hereon we will not make the distinction between a set of rules $R$ and the induced binary relation $\underset{R}{\longrightarrow}$.

**Definition 1.12 (Diamond Property)** *Let $R$ be a binary relation on a set $X$. Then $R$ has the diamond property (notation $R \models \diamond$ ) if*

$$\forall\, t,\ t_1,\ t_2 \in X\ [t \longrightarrow t_1 \wedge t \longrightarrow t_2 \ \Longrightarrow\ \exists\, t_3\ [t_1 \longrightarrow t_3 \wedge t_2 \longrightarrow t_3\ ]]$$

See diagram below:

$$
\begin{tikzcd}
 & t \arrow[dl] \arrow[dr] & \\
t_1 \arrow[dr] & & t_2 \arrow[dl] \\
 & t_3 &
\end{tikzcd}
$$

Remark: $\mathrm{R} \models \diamond \implies \mathrm{R}^* \models \diamond$.
Remark: $\mathrm{R} \models \diamond \implies \mathrm{R} \models \mathrm{CR}$.

**Definition 1.13 (Underlining)** *Let $R$ be a binary relation in $X$, define $\underline{R}$ and $\underline{X}$ as follows*

- *$\underline{R}$ is the binary relation in $\underline{X}$, obtained by underlining all the leftmost function symbols in the left-hand-side of the reduction rules in $R$.*
- *$\underline{X}$ is the set containing all terms in $X$, plus terms with some function symbols underlined.*

There are operations that allow to go from the structure (X,R) to ($\underline{\mathrm{X}}$,$\underline{\mathrm{R}}$) and vice-versa. One can convert a term $t$ in X to $t'$ in $\underline{\mathrm{X}}$ by possibly underlining some function symbols (lifting). Conversely, a term $t'$ in $\underline{\mathrm{X}}$ can be converted to $t$ in X by erasing all underlinings ( *i.e.* $t = \mid t' \mid$) . More formally:

**Lemma 1.14** *(i)*

$$
\begin{tikzcd}
t' \arrow[r, dashed, twoheadrightarrow] \arrow[d, "||"'] & t'_1 \arrow[d, dashed, "||"] & t', t'_1 \in \underline{X} \\
t \arrow[r, twoheadrightarrow] & t_1 & t, t_1 \in X
\end{tikzcd}
$$

*(i)*

$$
\begin{tikzcd}
t' \arrow[r, twoheadrightarrow] \arrow[d, "||"'] & t'_1 \arrow[d, "||"] & t', t'_1 \in \underline{X} \\
t \arrow[r, dashed, twoheadrightarrow] & t_1 & t, t_1 \in X
\end{tikzcd}
$$

**Definition 1.15 (Development with respect to $\mathcal{F}$)** *Given a term $t \in X$, and a set, $\mathcal{F}$, of redexex occurrence in $t$, let $t' \in \underline{X}$ be the term obtained by underlining the redexex in $\mathcal{F}$, then the reduction sequence $\sigma : t' \to t'_1 \to \ldots \to t'_n$ in $\underline{R}$ is called a development of $t$ with respect to $\mathcal{F}$ . A development of a term $t$ is a development of $t$ with respect to the set, $\mathcal{F}$, of all redex occurrences in $t$.*

Informally the previous definition says that a *development* of a term $t$ is a reduction in which only "old" redexex (*i.e.* , redexex already present in $t$), are rewritten.

**Definition 1.16 (Complete Development with respect to $\mathcal{F}$)** *Given a term $t \in X$, and a set, $\mathcal{F}$, of redexex occurrence in $t$, let $t' \in \underline{X}$ be the term obtained by underlining the redexex in $\mathcal{F}$, then the reduction sequence $\sigma : t' \to t'_1 \to \ldots \to t'_n$ in $\underline{R}$ is called a complete development of $t$, with respect to $\mathcal{F}$, if $t'_n$ does not contain any more underline. A complete development of a term $t$ is a complete development of $t$ with respect to set, $\mathcal{F}$, of all redex occurrences in $t$.*

The notions of "developments" and "complete developments" are very important because as we will see, they will allow us to prove confluence for both Regular TRS's and the $\lambda$-calculus.

## 2 Confluence for Regular TRS's

The proof of CR for Regular TRS's follows the steps below:

(i) R Regular $\Longrightarrow$ $\underline{\mathrm{R}}$ is Regular (lemma 2.1)

(ii) $\underline{\mathrm{R}} \models$ WCR ( lemma 2.1 and lemma 2.3)

(iii) $\underline{\mathrm{R}} \models$ SN (lemma 2.4)

(iv) (ii $\wedge$ iii) $\Longrightarrow$ $\underline{\mathrm{R}} \models$ CR (by Newman's lemma)

(v) R $\models$ CR (because $\mathrm{R}^* = \underline{\mathrm{R}}^*$)

The main point to grasp here is that in order to show that a reduction relation $R$ is CR, we define a new reduction relation $\underline{R}$, for which it's easier to show that is CR. We reduce the problem to something more tractable and the translation between the two different problems is given by showing that the two reduction relations have the same transitive closure. Therefore, once proved that $\underline{\mathrm{R}}$ is CR, it follows that $R$ is CR also.

**Lemma 2.1** *$R$ Regular $\Longrightarrow$ $\underline{R}$ Regular.*

*Proof:* Left to the reader.

■

**Fact 2.2** *Given a Regular TRS (X,R), $\forall\, t \in X$, $t \xrightarrow{\alpha} t_1 \wedge t \xrightarrow{\beta} t_2$ then the ($\alpha$) $\beta$-reduction does not modify the ($\beta$) $\alpha$-redex.*

At the first look it seems that the above is only due to non-overlapping patterns. Instead also non-left linearity can cause problems. As an example, given the rules

$$\begin{aligned} \mathsf{D}\, x\, x &\longrightarrow x \\ \mathsf{I}\, x &\longrightarrow x \end{aligned}$$

consider the term

$$\underbrace{\mathsf{D}\ \underbrace{(\mathsf{I}\ x)}_{\alpha}\ (\mathsf{I}\ x)}_{\beta}$$

the $\alpha$-reduction modify the $\beta$-redex.

**Lemma 2.3** *Given a Regular TRS (X,R), R is WCR.*

*Proof:* We want to show the following:

$$\forall\, t \in X,\ t \xrightarrow{\alpha} t_1\ \ t \xrightarrow{\beta} t_2 \implies \exists\, t_3,\ t_1 \twoheadrightarrow t_3 \quad \text{and } t_1 \twoheadrightarrow t_3$$

We do the proof by case analysis.

*Case* 1: **Redexes $\alpha$ and $\beta$ are disjoint**
Trivial. See diagram below:

$$\begin{tikzcd}
 & t \arrow[dl, "\alpha"'] \arrow[dr, "\beta"] & \\
t_1 \arrow[dr, "\beta"'] & & t_2 \arrow[dl, "\alpha"] \\
 & t_3 &
\end{tikzcd}$$

*Case* 2: Without loss of generality assume that $\alpha$ **is nested inside** $\beta$
Since R is regular, by the previous fact, only two cases are possible

(2.1) $\beta$-reduction destroys the $\alpha$-redex;

(2.2) $\beta$-reduction duplicates the $\alpha$-redex;

We consider the two cases above separately.

*SubCase* 2.1: **$\beta$-reduction destroys the $\alpha$-redex**
This means that

$$\beta \subseteq t_1$$

Suppose by contradiction that the above is not true, then the $\alpha$-reduction must have erased $\beta$. The only way this can happen is if $\beta \subseteq \alpha$. In conclusion:

$$\alpha \subseteq \beta \wedge \beta \subseteq \alpha \implies \alpha = \beta$$

We reached a contradiction, since R is not ambigous.

Therefore,

$$\begin{tikzcd}
t \arrow[r, "\beta"] \arrow[d, "\alpha"'] & t_2 \equiv t_3 \\
t_1 \arrow[ur, "\beta"'] &
\end{tikzcd}$$

*SubCase* 2.2: **$\beta$-reduction duplicates $\alpha$**

For the same reasons as before $\beta \subseteq t_1$, therefore

$$
\begin{tikzcd}
t \arrow[r, "\beta"] \arrow[dd, "\alpha"'] & t_2 \arrow[d, "\alpha"] \\
& \vdots \arrow[d, "\alpha"] \\
t_1 \arrow[r, "\beta"] & t_3
\end{tikzcd}
$$

■

**Lemma 2.4** *For any TRS (X,R), $\underline{R}$ is SN.*

*Proof:* The proof strategy is similar to the one given in ithe next section. ■

With these lemmas and definitions in our pocket, we can now state and prove the main theorem of this discussion.

**Theorem 2.5** *Given a Regular TRS (X,R), R is CR.*

*Proof:* Left to the reader. ■

## 3 The $\lambda$-calculus

Hereon $\Lambda$ indicates the set of $\lambda$-terms and $\beta$ indicates a reduction relation on $\Lambda$.

**Fact 3.1** $\beta \not\models \diamond$.

For an example, consider:

$$
\begin{tikzcd}
(_\alpha\, (\lambda\, x.x\, x)(_\beta\, II)) \arrow[r, "\alpha"] \arrow[dd, "\beta"'] & (_{\beta_0}\, II)\, (_{\beta_1}\, II) \arrow[d, "\beta_0"] \\
& I\, (_{\beta_1}\, II) \arrow[d, "\beta_1"] \\
(_\alpha\, (\lambda\, x.x\, x)\, I \arrow[r, "\alpha"] & II
\end{tikzcd}
$$

In the above example the $\alpha$-reduction duplicates the $\beta$-redex, the two copies are named $\beta_0$ and $\beta_1$ respectively. For this reason the $\lambda$-calculus has the so called ***duplicative*** property. This causes many issues regarding efficient implementations (say something about Vinod thesis).

Proof strategy of CR for $\lambda$-calculus:

(i) Define a new type of reduction relation, $\beta$'

(ii) $\beta' \models$ CR

(iii) $\beta^* = \beta'^*$.

By remark 1.3 it follows that $\beta^* \models \diamond$ and then by remark 1.3 we have proven that $\beta$ is CR.

### 3.1 Marked $\lambda$-calculus ( $\Lambda$')

In order to formalize the ideas of development and complete development, we introduce the new calculus $\Lambda$'. The $\Lambda$' terms are given by the following production:

$$E = x \mid \lambda\, x.E \mid E\ E \mid (\underline{\lambda}\, x.E)\ E$$

The rules of $\Lambda$' are:

$$\begin{aligned} \beta_0 :&\quad (\underline{\lambda}\, x.M)\ N \longrightarrow M\ [N/x] \\ \beta :&\quad (\lambda\, x.M)\ N \longrightarrow M\ [N/x] \end{aligned}$$

Notice that we do not underly arbitrary $\lambda$'s, only the ones that constitute the operator part of a redex. Thus, given the well-know term $(\lambda\, x.x\ x)\ (\lambda\, x.x\ x)$, you can certainly underline the first $\lambda$, obtaining $(\underline{\lambda}\, x.x\ x)\ (\lambda\, x.x\ x)$. However, you should convince yourself that $(\lambda\, x.x\ x)\ (\underline{\lambda}\, x.x\ x)$ is not a term in $\Lambda$'.

**Lemma 3.2** *Let $M \in \Lambda$ and $\mathcal{F}$ a set of redexex occurrences in M, then $\sigma$ is a development of M relative to $\mathcal{F}$ iff the lifted reduction $\sigma$', starting with $\underline{M}$, is a $\beta_0$-reduction. Where $\underline{M}$ is M with all the redexex in $\mathcal{F}$ underlined,*

**Definition 3.3** *Let $M \in \Lambda$ and $\mathcal{F}$ a set of redexex occurrences in M, then $\sigma\ : M \longrightarrow M_1 \cdots M_n$ is a complete development of M relative to $\mathcal{F}$ iff the lifted reduction $\sigma' : \underline{M} \longrightarrow \underline{M_1} \cdots \underline{M_n}$, is a $\beta_0$-reduction and $\underline{M_1}$ is in normal form.*

As an example, consider:

$$\begin{tikzcd}
({}_\alpha\ (\underline{\lambda}\, x.x\ x)({}_\beta\ \underline{I}(I\ a))) \arrow[r, "\alpha"] \arrow[dd, "\beta"'] & ({}_{\beta_0}\ \underline{I}\ (I\ a))\ ({}_{\beta_1}\ \underline{I}\ (I\ a)) \arrow[d, "\beta_0"'] \\
 & (I\ a)\ ({}_{\beta_1}\ \underline{I}\ (I\ a)) \arrow[d, "\beta_1"'] \\
({}_\alpha\ (\underline{\lambda}\, x.x\ x)\ (I\ a) \arrow[r, "\alpha"] & (I\ a)\ (I\ a)
\end{tikzcd}$$

### 3.2 Confluence for $\lambda$-calculus

**Lemma 3.4 (Substitution lemma)** *If $x \not\equiv y$ and $x \notin FV(L)$, then*

$$M\,[N/x][L/y] \equiv M\,[L/y][N\,[L/y]/x]$$

*Proof:*[By structural induction]

*Case* 1: **M is a variable**

*SubCase* 1.1: $\mathbf{M} \equiv \mathbf{x}$
Perform the substitution in both sides and you obtain

$$N\,[L/y] \equiv N\,[L/y]$$

*SubCase* 1.2: $\mathbf{M} \equiv \mathbf{y}$
Perform the substitution in both sides and you obtain

$$L \equiv L\,[N\,[L/y]/x] \equiv L \quad x \notin FV(L)$$

*SubCase* 1.3: $\mathbf{M} \equiv z \not\equiv x \not\equiv y$
In both sides we obtain z

*Case* 2: $\mathbf{M} \equiv M_1\,M_2$
Follows directly from induction hypothesis

*Case* 3: $\mathbf{M} \equiv \lambda\, z.M_1$

$$\begin{array}{rcll}(\lambda\, z.M_1)\,[N/x][L/y] & \equiv & \lambda\, z.(M_1\,[N/x][L/y]) & \text{by definition of substitution}\\ & \equiv & \lambda\, z.(M_1\,[L/y][N\,[L/y]/x]) & \text{by induction hyphothesis}\\ & \equiv & (\lambda\, z.M_1)\,[L/y][N\,[L/y]/x] & \text{by definition of substitution}\end{array}$$

■

**Lemma 3.5** $\beta_0 \models$ *WCR.*

*Proof:* Let $\rho_1$ and $\rho_2$ be the two redexex contracted, we will do the proof on case analysis on the relative position of $\rho_1$ and $\rho_2$.

*Case* 1: **$\rho_1$ and $\rho_2$ are disjoint**
Trivial

*Case* 2: Without loss of generality assume that $\rho_1 \subseteq \rho_2$
Assume that $\rho_1 \equiv (\underline{\lambda}\, x.P)\,Q$ and $\rho_2 \equiv (\underline{\lambda}\, x.M)\,N$.

*SubCase* 2.1: $\rho_1 \subseteq$ **M.**

Follows from Substitution lemma.

*SubCase* 2.2: $\rho_1 \subseteq$ **N.**

$$\begin{tikzcd}
({}_{\rho_2}\,(\underline{\lambda}\, x.\cdots x\cdots x)(\cdots({}_{\rho_1}\,(\underline{\lambda}\, y.P)Q)\cdots) \arrow[r, "\rho_2"] \arrow[dd, "\rho_1"'] & \cdots({}_{\rho_1}\,(\underline{\lambda}\, y.P)\,Q)\cdots({}_{\rho_1}\,(\underline{\lambda}\, y.P)\,Q)\cdots \arrow[d, "\rho_1"'] \\
 & \vdots \arrow[d, "\rho_1"'] \\
({}_{\rho_2}\,(\underline{\lambda}\, x.\cdots x\cdots x)(\cdots P\,[Q/y]\cdots)) \arrow[r, "\rho_2"] & \cdots P\,[Q/y]\cdots P\,[Q/y]\cdots
\end{tikzcd}$$

■

The main technique to prove that a binary relation in a set $X$ is SN, is to show that that binary relation *well-order* $X$, that is, each chain in $X$ has a minimum element. In our case we are interested in showing that $\beta_0$ is SN, thus we proceed as follows:

- Assign a *weight* to each $M \in \Lambda'$, call the term so obtained $\mid M \mid$
- show:

$$M \longrightarrow N \implies \mid N \mid < \mid M \mid$$

that is, the "weight" of a term is decreasing as we reduce it.

**Definition 3.6 (Weighting)** *Given $M$ in $\Lambda$ associate a positive integer to each variable occurrence in $M$.*

We thus obtain a new calculus, $\Lambda^*$, that has the usual inductive definition with the variables ranging over $x^0 \cdots x^n$. The definition of reduction on $\Lambda^*$ ($\beta_0^*$) carries over in the usual way.

**Definition 3.7 (Weight)** *Let $M$ in $\Lambda^*$ define $\mid M \mid$ as the sum of the weights occurring in $M$.*

**Definition 3.8** *Let $M$ in $\Lambda^*$, then $M$ has decreasing weight property (dwp) if for every $\beta_0^*$-redex $(\underline{\lambda} x.P)Q$ in $M$ :*

$$\forall x \in P,\ \mid x \mid > \mid Q \mid$$

Example: $(\underline{\lambda} x.x^6x^7)(\underline{\lambda} x.x^2x^3)$ has the *dwp*, while $(\underline{\lambda} x.x^4x^7)(\underline{\lambda} x.x^2x^3)$ does not.

**Lemma 3.9** *For all $M$ in $\Lambda^*$, there exists an initial weight assignement so that $M$ has decreasing weight property.*

*Proof:* Start enumerating all variables occurrences in $M$ from right to left, and assign to, let's the $m^{th}$ variable occurence the weight $2^m$. Since

$$2^m \;>\; 2^{m-1} + 2^{m-2} + \cdots 2 + 1$$

$M$ has the *dwp*.

■

**Lemma 3.10** *If $M \longrightarrow N$, and $M$ has dwp then*

$$|N| < |M|$$

*Proof:* Let $M$ be $\cdots(\underline{\lambda}x.P)Q\cdots$

*Case* 1: $x \notin P$
Then $Q$ vanishes

*Case* 2: $x \in P$
The weight must decrease because the weight of the substituted expression, *i.e.* $|Q|$, is less than every $x$.

■

**Lemma 3.11** *Let $M \longrightarrow N$, then if $M$ has dwp so does $N$.*

*Proof:* Suppose $M \xrightarrow{R_0} N$, where $R_0 \equiv (\underline{\lambda}x.P_0)Q_0$. Examine the effect of $R_0$-reduction on some other redex $R_1 \equiv (\underline{\lambda}y.P_1)Q_1$ in M. We will do the analysis on the relative positions of $R_0$ and $R_1$.

*Case* 1: $R_0 \cap R_1 = \emptyset$
$R_0$-reduction does not affect $R_1$

*Case* 2: $R_1 \subseteq R_0$

*SubCase* 2.1: **$R_1$ is inside the rator $\underline{\lambda}x.P_0$**

$$R_0 \equiv (\underline{\lambda}x.\cdots((\underline{\lambda}y.P_1)Q_1)\cdots)Q_0.$$

By the dwp of M,

$$\forall\, y \in P_1,\ |y| > |Q_1|$$

and, by the fact that $y \notin \mathrm{FV}(Q_0)$,

$$\forall\, y \in P_1\,[Q_0/x],\ |y| > |Q_1|$$

And,

$$\forall\, x \in R_0,\ |x| > |Q_0|$$

then

$$|Q_1| > |Q_1\,[Q_0/x]|$$

In conclusion,

$$\forall\, y \in P_1\,[Q_0/x],\ |y| > |Q_1\,[Q_0/x]|$$

*SubCase* 2.2: **$R_1$ is inside the rand $Q_0$**

$$R_0 \equiv (\underline{\lambda} x.P_0)(\cdots R_1 \cdots)$$

$R_0$-reduction does not modify $R_1$ (may just copy it or destroy it)

*Case* 3: $R_0 \subseteq R_1$

*SubCase* 3.1: **$R_0$ is inside the rator of $R_1$**

$$R_1 \equiv (\underline{\lambda} y. \cdots ((\underline{\lambda} x.P_0)Q_0) \cdots)Q_1$$

The weight of $y$ is not affected by $R_0$-reduction.

*SubCase* 3.2: **$R_0$ is inside the rand of $R_1$**

$$R_1 \equiv (\underline{\lambda} y.P_1)(\cdots ((\underline{\lambda} x.P_0)Q_0) \cdots)$$

The weight of $Q_1$ after $R_0$-reduction decreases.

■

From the previous lemma we can infer,

**Lemma 3.12** *$\beta_0 \models SN$.*

**Corollary 3.13** *$\beta_0 \models CR$ (by Newman's lemma).*

**Theorem 3.14 (Finite Development)** *Let $M \in \Lambda$ and $\mathcal{F} \subseteq M$*

*(i) All developments of $M$ related to $\mathcal{F}$ are finite;*

*(ii) All complete developments of $M$ related to $\mathcal{F}$ end up with the same term.*

*Proof:*
(i) follows from lemma 3.12
(ii) follows from lemma 3.13

■

We can now define the new reduction relation,

**Definition 3.15 (Parallel reduction)** *$M \underset{1}{\twoheadrightarrow} N$, if $N$ is the result of a complete development of $M$ with respect to some $\mathcal{F}$.*

Exercise:
Let M $\equiv (\lambda\ x.x\ x)(I\ I)$. Then it is a good exercise to see what M parallel reduce to. In particular, does M $\underset{1}{\twoheadrightarrow} I\ (I\ I)$ ?.

**Theorem 3.16** $\underset{1}{\longrightarrow} \models \diamond$.

*Proof:*

$$\begin{tikzcd}
t \arrow[r, "\mathcal{F}_2"] \arrow[d, "\mathcal{F}_1"'] \arrow[dr, "\mathcal{F}_1 \cup \mathcal{F}_2"] & t_2 \arrow[d, dashed, "\mathcal{F}_1"] \\
t_1 \arrow[r, dashed, "\mathcal{F}_2"'] & t_3
\end{tikzcd}$$

■

# 4 Appendix

**Definition 4.1 (Least element)** *The least element of a partially order set $(X, \leq)$ is a $y \in X$ such that*

$$\forall\, x \in X,\; y \;\leq\; x$$

Note that if the least element exists, it is unique, so one should talk of *the* least element of $X$.

**Definition 4.2 (Well-ordered set)** *A partially ordered set $(X,R)$ is well-ordered iff each non-empty subset has a least element.*

Examples: The set $\{1, 2, 3, \cdots\}$ is well-ordered. The set $\{\cdots, 3, 2, 1\}$ is not well-ordered, because it has no first element.

Remark: One consequence of this definition is that every well-ordered set $X$ is totally ordered. Let $x, y \in X$, then $\{x, y\}$ is a non-empty subset of $X$ and has therefore a least element. If the least element is $x$, then $x\ R\ y$, otherwise, $y\ R\ x$.

The reason of the interest in well-ordered sets, lies in the fact that we can prove properties of their elements using a process similar to **mathematical induction.**

**Definition 4.3 (Transfinite induction)** *Given $S$, a subset of a well-ordered set $X$,*

$$[\forall\, x \;\in\; X, s(x) \;\subseteq\; S \;\Longrightarrow\; x \;\in\; S] \;\;\Longrightarrow\;\; S = X$$

**Definition 4.4** *Let the set $S$ be a subset of the partial order $X$ then*

$$S \sqsubseteq y,\;\; \textit{if} \forall\, x \in S\;\; [x \sqsubseteq y]$$

**Definition 4.5 (Least Upper Bound)** *Let the set $X$ be partially ordered by $R$, $z$ is the least upper bound of a subset $S$ of $X$ (denoted as $\bigsqcup S$) iff:*

1. $S \sqsubseteq z$.
2. $\forall x \in X,\; S \sqsubseteq x \Longrightarrow z \sqsubseteq x$

Example:

$$\begin{tikzcd} a & c \\ b \arrow[u] \arrow[ur] & d \arrow[u] \arrow[ul] \end{tikzcd}$$

$\{$a, b$\}$ = a

$\{$a, c$\}$ has no upper bound,so no lub

$\{$b, d$\}$ has two upper bounds, a and c but not lub

**Definition 4.6 (Directed set)** *Let the set $X$ be partially ordered by $R$. A nonempty set $S$ is directed iff:*

$$\forall x_1, x_2 \in S,\; \exists x_3 \;\in S \textit{ such that } x_1 \sqsubseteq x_3 \textit{ and } x_2 \sqsubseteq x_3.$$

To give an analogy, directed sets are like the Cauchy sequences in analysis.

**Example 1:** {a, b, d} of the previous example is directed.
**Example 2:** Given a set $X$, define $S$ to be the set of all finite subsets of X:

$$S = \{F \subseteq X : \ \mid F \mid < \omega\}$$

$S$ is partially ordered by subset inclusion and is directed:

$$\forall F_1, F_2 \in S, \ F_1 \cup F_2 \in S \text{ is an upper bound.}$$

**Definition 4.7 (Chain)** *For a partially ordered set $X$, a subset $S$ of $X$ is a chain iff $S$ is nonempty, and $\forall x_1, \ x_2 \in S, \ x_1 \sqsubseteq x_2$ or $x_2 \sqsubseteq x_1$.*

Directed sets are more general than chains, for example, Example 1 is not a chain because b and d are incomparable. More significant, in Example 2, if X is uncountable then no chain of elements in $S$ has the same lub as S.

**Definition 4.8** *A partially ordered set $(X, R)$ is a complete partial order (cpo) iff:*

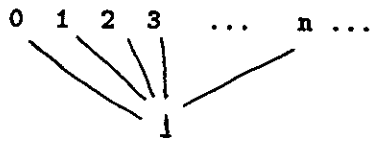
1. $\exists \bot \in X$ *s.t.* $\bot \sqsubseteq X$.
2. $\forall$ *directed* $S \subseteq X, \ \bigsqcup S \in X$ *exists.*

Despite the technical distinction between chains and directed sets the difference won't bother us. For example, $cpo_{chain} = cpo_{dir}$, since if every chain has a lub then every directed set has a lub.

**Example 1:** $(P(X), \subseteq)$ is a cpo:

1. there exists the least element $\bot = \emptyset$;
2. $\forall$ directed $S \subseteq P(X), \ \bigsqcup S = \bigcup S \in P(X)$.

**Example 2:** $(N_\bot, \sqsubseteq)$ is a cpo with $\bot \sqsubseteq N$.




N.B. $N_\bot$ captures the intuition that a program running forever ($\bot$) is less defined than a program converging to some result.

**Definition 4.9** *A function $f : X \to Y$, where $X$ and $Y$ are partially ordered sets, is monotone iff $\forall c_1, c_2 \in X, \ c_1 \sqsubseteq_X c_2 \implies f(c_1) \sqsubseteq_Y f(c_2)$. The function $f$ is continuous iff $f$ is monotone and $\forall$ directed $S \subseteq X, \ f(\bigsqcup S) = \bigsqcup\{f(x) : x \in S\}$.*

Note monotone functions need not be continuous, e.g., let $A = \{0, 1, ..\} \cup \{\omega, \omega + 1\}$ and $f : A \to A$ where $f(n) = n \; \forall n \in N$, $f(\omega) = f(\omega + 1) = \omega + 1$. $f$ not continuous can be seen from the fact that it does not commute with taking the limit:

$$f(\bigsqcup S) = f(\omega) = \omega + 1 \neq \omega = \bigsqcup S = \bigsqcup f(S)$$

We denote the set of all continuous functions from $X$ to $Y$ by $X \to_c Y$, and the set of all monotone functions by $X \to_m Y$.

**Lemma 4.10** *If $X, Y$ are cpo's, then so are $X \to_c Y$, $X \to_m Y$ with the order $f \sqsubseteq g$ iff $\forall x \in X,\; f(x) \sqsubseteq g(x)$.*

*Proof:*

- $\sqsubseteq$ is a partial order.
- There is a least element defined as:

  $$\perp_{X \to_m Y}(x) = \perp_Y$$

  i.e., $\perp_{X \to_m Y}$ is everywhere less than or equal to every function:

  $$\forall x \in X,\; z \in X \to_m Y,\; \perp_{X \to_m Y}(x) = \perp_Y \sqsubseteq g(x)$$

- Let $F \subseteq (X \to_m Y)$ be a directed set. Define the function $g$ by

  $$g(x) = \bigsqcup \{f(x) : f \in F\}$$

  Since $F$ is a directed set :

  $$\forall f_1, f_2 \;\exists f_3 \in F \; s.t. \; f_1 \sqsubseteq f_3 \text{ and } f_2 \sqsubseteq f_3$$

  This implies:

  $$\forall x \in X \; f_1(x) \sqsubseteq f_3(x) \text{ and } f_2(x) \sqsubseteq f_3(x)$$

  We conclude that $\{f(x) : f \in F\}$ forms a directed set for each $x$. Thus, it has a lub, so $g(x)$ is well defined and clearly $g$ is a upper bound of $F$:

  $$\forall f_i \in F,\; \forall x \in X \; f_i(x) \sqsubseteq g(x) \Longrightarrow F \sqsubseteq g$$

  Moreover, it is the least upper bound, for if there was a smaller upper bound, there would exist some $x$ such that $g(x)$ were not the lub of the $\{f(x) : f \in F\}$.

- we have to prove now that $g$ is monotone. Suppose $x \sqsubseteq x'$, then by monotonicity of $f$:

  $$\forall f \in F,\; f(x) \sqsubseteq f(x')$$

  and this implies:

  $$\bigsqcup \{f(x) : f \in F\} \sqsubseteq \bigsqcup \{f(x') : f \in F\}$$

  Thus:

  $$g(x) = \bigsqcup \{f(x)\} \sqsubseteq \bigsqcup \{f(x')\} = g(x').$$

We have thus shown that a set of monotone functions is a cpo. We state without proof that $X \to_c Y$ is also a cpo.

**Theorem 4.11 (Tarski-Knaster)** *Let $X$ be a cpo and let $f : X \to_m X$. Then $f$ has a least fixed point.*

*Proof:*

Define the sequence of values:

$$x_0 = \perp$$
$$x_\alpha = \bigsqcup\{f(x_\beta) : \beta < \alpha\}$$

where $\alpha$ range over the ordinals.

For any ordinal $\alpha$ by transfinite induction we have $x_\alpha \sqsubseteq f(x_\alpha)$. This implies :

$$\forall\beta \le \alpha,\ x_\beta \sqsubseteq x_\alpha$$

So, $\{f(x_\beta) : \beta < \alpha\}$ is a directed set in $X$. Consider the directed set $x_0, x_1, \ldots x_\kappa$, such that $\kappa > \mid X \mid$. By pigeonholing there must exist $\kappa_1, \kappa_2$ s.t. $\kappa_1 < \kappa_2 \le \kappa$, and $x_{\kappa_1} = x_{\kappa_2}$. But:

$$f(x_{\kappa_1}) \sqsubseteq x_{\kappa_2} \Longrightarrow f(x_{\kappa_1}) \sqsubseteq x_{\kappa_1}$$

Moreover we know $x_{\kappa_1} \sqsubseteq f(x_{\kappa_1})$, it follows that $f(x_{\kappa_1}) = x_{\kappa_1}$, hence $x_{\kappa_1}$ is a fixed point of $f$.

To see that $x_{\kappa_1}$ is the least fixed point, suppose $y$ is a fixed point of $f$. By monotonicity of $f$:

$$\perp \sqsubseteq y$$
$$f(\perp) \sqsubseteq f(y) = y$$
$$f^2(\perp) \sqsubseteq f^2(y) = y$$
$$\vdots$$

by induction:

$$x_{\kappa_1} \sqsubseteq y.$$

■

**Lemma 4.12** $\mu : (X \to X) \to X$ *is monotone.*

*Proof:* Suppose $f \sqsubseteq g$. By monotonicity:

$$x_0^f \sqsubseteq x_0^g$$
$$x_1^f \sqsubseteq x_1^g$$
$$\vdots$$

by induction:

$$\mu f \sqsubseteq \mu g$$

■

# 5 Notation

| | |
|---|---|
| $\twoheadrightarrow_R$ | reduction relation induced by R |
| $\overset{+}{\twoheadrightarrow}$ | 1 or more steps reduction |
| $\overset{n}{\twoheadrightarrow}$ | n steps reduction |
| $\xrightarrow{\alpha}$ | reduction of the $\alpha$-redex |
| $\alpha \subseteq t$ | $\alpha$ is a subterm of t |

# Contents


**1 Basic Definition and Properties** **1**

1.1 Order . . . . . . . . . . 1

1.2 Application of Noetherian Induction . . . . . . . . . . 3

1.3 Reduction Properties . . . . . . . . . . 4

**2 Confluence for Regular TRS's** **6**

**3 The $\lambda$-calculus** **8**

3.1 Marked $\lambda$-calculus ( $\Lambda$') . . . . . . . . . . 9

3.2 Confluence for $\lambda$-calculus . . . . . . . . . . 10

**4 Appendix** **14**

**5 Notation** **18**