The Confluence of the λ-calculus

Arvind
Laboratory for Computer Science
M.I.T.

December 4, 2002

http://www.csg.lcs.mit.edu/6.827

Confluence *aka Church-Rosser Property*

A reduction system R is said to be *confluent (CR)*,
if $t \twoheadrightarrow t_1$ and $t \twoheadrightarrow t_2$ then there exits a t_3 such that
$t_1 \twoheadrightarrow t_3$ and $t_2 \twoheadrightarrow t_3$.

$$\begin{tikzcd} t \arrow[r, two heads] \arrow[d, two heads] & t_1 \arrow[d, two heads] \\ t_2 \arrow[r, two heads] & t_3 \end{tikzcd}$$

Fact: In a confluent system, if a term has a normal form then it is *unique* .

Theorem: The λ–calculus is confluent.

Theorem: An orthogonal TRS is confluent.

The Diamond Property

A reduction system R is said to have the *diamond property* , if t --> t_1 and t --> t_2 then there exits a t_3 such that t_1 --> t_3 and t_2 --> t_3.

$$\begin{tikzcd} t \arrow[r] \arrow[d] & t_1 \arrow[d] \\ t_2 \arrow[r] & t_3 \end{tikzcd}$$

Theorem: If R has the diamond property then R is confluent.

Fact: The λ-calculus does not have the diamond property.

Weak Confluence

A reduction system R is said to be *weakly confluent (WCR)*, if t $\rightarrow$ t_1 and t $\rightarrow$ t_2 then there exits a t_3 such that $t_1 \twoheadrightarrow t_3$ and $t_2 \twoheadrightarrow t_3$.

$$\begin{tikzcd} t \arrow[r] \arrow[d] & t_1 \arrow[d, two heads] \\ t_2 \arrow[r, two heads] & t_3 \end{tikzcd}$$

In a WCR system one step divergence can be contained!

Theorem: If R is CR then R is also WCR.

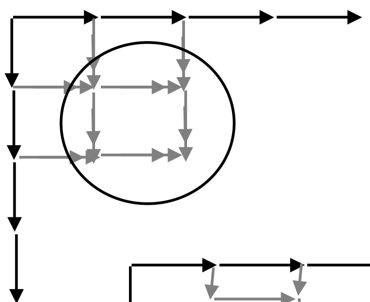
Theorem: If R is WCR then $\underline{R}$ is also WCR.

WCR does not imply CR

Example:

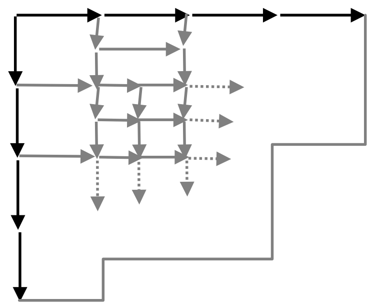
F(x) $\rightarrow$ G(x)
F(x) $\rightarrow$ 1
G(x) $\rightarrow$ F(x)
G(x) $\rightarrow$ 0

Why WCR does not imply CR



Suppose R is WCR

Completing this diagram looks like proving the CR theorem again!



The diagram may not complete!

There will be no problem if all the reduction paths were finite

Strongly Normalizing Systems

Let (Σ , R) be a TRS and *t* be a term

t is in *normal form* if it cannot be reduced any further.

Term *t* is *strongly normalizing (SN)* if every reduction sequence starting from *t* terminates eventually.

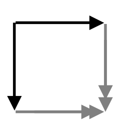
R is *strongly normalizing (SN)* if for all terms every reduction sequence terminates eventually.

R is *weakly normalizing (WN)* if for all terms there is some reduction sequence that terminates.

Neumann's Lemma

If a reduction system R is SN and WCR then R is CR.

How does it help us when an R is not SN ?



Only “old” redexes need to be performed to close the diagram

$\Rightarrow$define a new reduction system for doing just the“old” redexes.

Is such a system SN?

Underlining and Development

Underline some redexes in a term.

Development is a reduction of the term such that only underlined redexes are done.

Complete Development is a reduction sequence such that all the underlined redexes have been performed.

$(\underline{S}\ \ K\ \ x\ \ (\underline{K}\ \ y\ \ z))$

$\rightarrow (\underline{S}\ \ K\ \ x\ \ y)$

$\rightarrow K\ y\ \ (x\ \ y)$

$\rightarrow K\ \ (\underline{K}\ \ y\ \ z\ )\ \ (x\ \ (\underline{K}\ \ y\ \ z\ ))$

$\rightarrow K\ \ y\ \ (x\ \ (\underline{K}\ \ y\ \ z))$

$\rightarrow K\ y\ \ (x\ \ y)$

By underlining redexes we can distinguish between old and newly created redexes in a reduction sequence.

The Underlined λ-calculus

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

Reduction rules:

β:?(?λx.M?) A --> ?M?[A/x] *the λ-calculus*

$\underline{\beta}$: $(\underline{\lambda}x.M)\ A \text{-->} M[A/x]$ *the $\underline{\lambda}$-calculus*

$\beta' = \beta \cup \underline{\beta}$

Erasure:

? Er? $\underline{\lambda}$-term --> λ-term

Facts:

$$\begin{tikzcd} \underline{M} \arrow[r, two heads, "\beta'"] \arrow[d, "Er"'] & \underline{N} \arrow[d, "Er"] \\ M \arrow[r, two heads, "\beta"] & N \end{tikzcd} \qquad \begin{tikzcd} \underline{M} \arrow[r, two heads, "\underline{\beta}"] \arrow[d, "Er"'] & \underline{N} \arrow[d, "Er"] \\ M \arrow[r, two heads, "\beta"] & N????? \end{tikzcd}$$

Complete Development *An Example*

M = (λx.x x) (I (I a))   where I = (λx.x)

Underline some redexes

M = ($\underline{\lambda}$x.x x) ($\underline{I}$ (I a))

→ ($\underline{I}$ (I a)) ($\underline{I}$ (I a))
→ (I a) ($\underline{I}$ (I a))
→ (I a) (I a)

→ ($\underline{\lambda}$x.x x) (I a)
→ (I a) (I a)

Underlined Reduction Systems are SN

Theorem: For every reduction system R, $\underline{R}$ is strongly normalizing.

Proof strategy:
Assign a *weight* to each term M such that the weight decreases after each reduction.

M -->$_{\_}$ N ⇒?? |N| < |M|

where |M| represents the weight of M.

Thus, if

M -->$_{\_}$ M_1 --> M_2 --> ...

⇒ |M| > $|M_1|$ > $|M_2|$ > ...

⇒ since for all M, |M| > 0, the reduction terminates!

Decreasing weight property

Assigning Weights (The λ–calculus)

Associate a positive integer to each *variable occurrence* in M

| M |: sum of the weights occurring in M

$\| x^w \|$	=	w
$\| \lambda x.M \|$	=	$\| M \|$
$\| \underline{\lambda} x.M \|$	=	$\| M \|$
$\| M\ N \|$	=	$\| M \| + \| N \|$

Weights, like underlined λ, are carried through the reduction unchanged.

Decreasing Weight Property (dwp)

M has *decreasing weight property* if for every $\underline{\beta}$-redex $((\underline{\lambda} x.P)\ Q)$ in M, |x| > |Q| for each free occurrence of x in P

Examples

$$M_1 = (\underline{\lambda} x.\ x^6\ \ x^7)\ (\underline{\lambda} y.\ y^2\ \ y^3)$$

$$M_2 = (\underline{\lambda} x.\ x^4\ \ x^7)\ (\underline{\lambda} y.\ y^2\ \ y^3)$$

Initial Weight Assignment

Lemma: There exits an initial weight assignment for each M such that M has dwp.

Proof:

1. Assign the weight 2^m to the m^{th} variable occurrence from the right

$$|\leftarrow \quad m \quad \rightarrow|$$
$$M = \quad . . . x . . . . . .$$
$$\Rightarrow |x| = 2^m$$

2. M has the dwp since

$$2^n > 2^{n-1} + 2^{n-2} + \ldots + 1$$

Example:

$$(x\ y\ ((\lambda z.z)\ \ (x\ x)))$$

Reduction Decreases the Weight

of a term with dwp

Lemma: If M has dwp and M --> N then |N| < |M|

Proof:

Suppose (($\underline{\lambda}$x.P) Q) is the redex that is reduced when M --> N.

Cases

(i) x is not in FV(P) :

(ii) x is in FV(P) :

dwp is Preserved Under Reduction

Lemma: If M --> N and M has dwp then so does N.

Proof: Suppose M --> N by doing the redex $R_0 \equiv (\lambda x.P_0)\ Q_0$. Examine the effect of R_0-reduction on some other redex $R_1 \equiv (\lambda y.P_1)\ Q_1$ in M.

Cases on relative position of R_0 and R_1

1. R_0 and R_1 are *disjoint*
2. R_1 is inside R_0 *(effect on subterms)*
3. R_0 is inside R_1 *(effect on the context)*

dwp is Preserved Under Reduction
continued-1

Suppose M --> N by doing the redex $R_0 \equiv (\underline{\lambda}x.P_0)\ Q_0$. Examine the effect of R_0-reduction on $R_1 \equiv (\underline{\lambda}y.P_1)\ Q_1$.

Case 2. R_1 is inside R_0 *(effect on subterms)*

2.1 R_1 is inside the rator, $\underline{\lambda}x.P_0$

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

2.2 R_1 is inside the rand, Q_0

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

dwp is Preserved Under Reduction
continued-2

Suppose M --> N by doing the redex $R_0 \equiv (\underline{\lambda}x.P_0)\ Q_0$.
Examine the effect of R_0-reduction on $R_1 \equiv (\underline{\lambda}y.P_1)\ Q_1$.

Case 3. R_0 is inside R_1 *(effect on the context)*

3.1 R_0 is inside the rator of R_1

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

3.2 R_0 is inside the rand of R_1

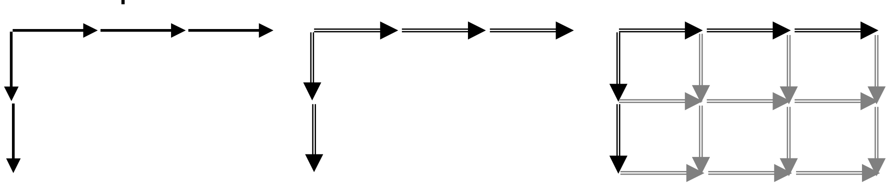
$$R_1 \equiv (\underline{\lambda}y.P_1)\ (...((\underline{\lambda}x.P_0)\ Q_0)...)$$

Proof Strategy for CR

Define a new type of reduction called complete developments (CD) using the underlined λ-calculus.

Prove the diamond property for CD reductions, i.e., show that CD is SN and CD is WCR.

The proof of confluence for the λ-calculus follows:



Each reduction can be viewed as a CD

Since CD reductions have the diamond property

λ- calculus is WCR

Suppose M --> M_1 by doing redex R_1 and M --> M_2 by doing redex R_2 .

We want to show that there exists an M_3 such that M_1 -->> M_3 and M_2 -->> M_3.

Cases on relative position of R_1 and R_2 in M.

1. R_1 and R_2 are *disjoint*

2. Without loss of generality assume R_1 is inside R_2
   2.1 R_1 is in the rator of R_2
   from the substitution lemma
   2.2 R_1 is in the rand of R_2

Substitution Lemma

If x is not equal to y and x is not in FV(L) then

M [N/x] [L/y] = M [L/y] [N[L/y]/x]

(λy.(λx.M) N) L

Finite Development Theorem

Suppose M is a λ-term and F is a set of redexes in M, then

1. All developments of M related to F are finite
2. All complete developments of M related to F end with the same term.

The proof follows from the fact that the $\underline{\lambda}$-calculus is SN and WCR

CD Reduction has the Diamond Property

$$\begin{tikzcd} M \arrow[r, two heads, "F_1"] \arrow[d, "F_2"'] & M_1 \arrow[d] \\ M_2 \arrow[r, two heads] & M_3 \end{tikzcd}$$

M_3 is a CD of M with respect to F_1 U F_2

$$\begin{tikzcd} M \arrow[r, Rightarrow] \arrow[d, Rightarrow] & M_1 \arrow[d, Rightarrow] \\ M_2 \arrow[r, Rightarrow] & M_3 \end{tikzcd}$$

Orthogonal TRS

- Confluence of orthogonal TRS's can be shown in the same way.

Orthogonal TRSs

A TRS is *Orthogonal* if it is:

1. *Left Linear:* has no multiple occurrences of a variable on the LHS of any rule, and
2. *Non Interfering:* patterns of rewrite rules are pairwise non-interfering

Theorem: An Orthogonal TRS is Confluent.

Orthogonal TRSs are CR

Proof outline:

1. R is orthogonal ⇒ $\underline{R}$ is orthogonal.

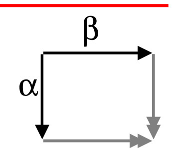
2. R is orthogonal ⇒?R is WCR ⇒?$\underline{R}$ is WCR.

3. $\underline{R}$ is SN

4. From 2. and 3. $\underline{R}$ is CR (Neumann's Lemma)

5. Transitive Closure of R = Transitive closure of $\underline{R}$
⇒ R is CR.

If R is orthogonal then R is WCR



Case 1: α and β are disjoint
α ?and β commute (trivially)
Case 2: α is a subexpression of β
(⇒ β cannot be a subexpression of α?

????*Case 2a: α?is reduced before β*
Since R is orthogonal, reducing α?
cannot affect β
????*Case 2b: β?is reduced before α*
????????β can destroy or duplicate α????????????????????