Chew's theorem revisited - uniquely normalizing property of nonlinear term rewriting systems -

Mizuhito Ogawa

NTT Basic Research Laboratories
3-9-11 Midori-cho Musashino-shi Tokyo 180 Japan
mizuhito@ntt-20.ntt.jp

Abstract. This paper gives a purely syntactical proof, based on proof normalization techniques, of an extension of Chew's theorem. The main theorem is that a weakly compatible TRS is uniquely normalizing. Roughly speaking, the weakly compatible condition allows possibly nonlinear TRSs to have nonroot overlapping rules that return the same results. This result implies the consistency of CL-pc which is an extension of the combinatory logic CL with parallel-if rules.

1 Introduction

The Church-Rosser (**CR**) property is one of the most important properties for term rewriting systems (TRSs). When a TRS is nonterminating, a well-known condition for **CR** is Rosen's theorem, which states that a left-linear weakly nonoverlapping TRS is **CR** - or, simply, that a left-linear nonoverlapping TRS is **CR**[11, 13]. A pair of reduction rules is said to be overlapping if their applications interfere with each other (i.e., they are unified at some nonvariable position), and a TRS is said to be nonoverlapping if none of its rules are overlapping (except that a same rule overlaps itself at the root). A TRS is said to be weakly nonoverlapping if applications of an overlapping pair of rules return the same result. Without the assumption of linearity, on the other hand, **CR** for a nonoverlapping TRS is not guaranteed for the following two reasons:

(i) A pair of nonoverlapping rules may overlap modulo equality. For instance, R_1 has a sequence $d(2,2) \to d(2,f(2)) \to d(f(2),f(2)) \to d(f(2),f^2(2)) \to \cdots$ s.t. $d(2,2)$, $d(f(2),f(2))$, $\cdots$ are reduced to 0, and $d(2,f(2))$, $d(f(2),f^2(2))$, $\cdots$ are reduced to 1. Because 0 and 1 are normal forms, R_1 is not **CR**[11].

(ii) A reducible expression (redex) for a nonlinear rule may not be recovered after some reduction destroys the identity of nonlinear variables. For instance, R_2 has a sequence $1 \to f(1) \to d(1,f(1)) \to d(f(1),f(1)) \to 0$. Thus, $1 \xrightarrow{*} 0$ and $1 \to f(1) \xrightarrow{*} f(0)$. Since 0 is a normal form and $f(0)$ simply diverges to $d(0,d(0,d(\cdots)))$, R_2 is not **CR**[3].

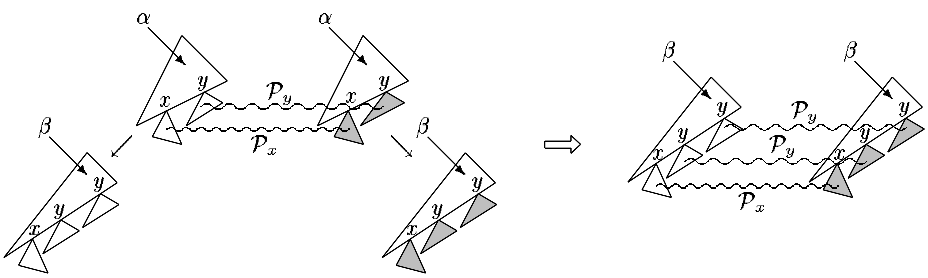
$$R_1 = \left\{ \begin{array}{ll} d(x,x) & \to 0 \\ d(x,f(x)) & \to 1 \\ 2 & \to f(2) \end{array} \right\} \qquad R_2 = \left\{ \begin{array}{ll} d(x,x) & \to 0 \\ f(x) & \to d(x,f(x)) \\ 1 & \to f(1) \end{array} \right\}$$

Chew and Klop have shown the sufficient condition for the uniquely normalizing (**UN**) property instead of **CR** - that is, a strongly nonoverlapping TRS is **UN**[3, 7]. A TRS is said to be strongly nonoverlapping if its linearization (i.e., the renaming of repeated variables with fresh individual variables) is nonoverlapping. Chew also states, more generally that a compatible TRS is **UN**[7]. In the compatible case, however, Chew's proof is hard to recognize, and its journal version has not yet been published[4].

Several trials to get a new proof have been reported[3, 15], and they show partial answers. Their main technique is to first transform a nonlinear TRS to a linear TRS (either unconditional or conditional) and use its weakly nonoverlapping property to prove it **CR**. Then, they translate the **CR** property of the linearized TRS to the **UN** property of the original nonlinear TRS[7, 8]. De Vrijer uses a similar technique to show that CL-pc (combinatory logic with parallel-if) is **UN**[15] and the key of his proof is the consistency check (i.e., $T \neq F$) accomplished by constructing a model of CL-pc. Finally the consistency shows that an application of $Czxx \to x$ may overlap modulo equality with an application of either $CTxy \to x$ or $CFxy \to y$, but not both.

$$\mathsf{CL\text{-}pc} = \mathsf{CL} \cup \left\{ \begin{array}{ll} CTxy & \to x \\ CFxy & \to y \\ Czxx & \to x \end{array} \right\} \qquad \text{where} \qquad \mathsf{CL} = \left\{ \begin{array}{ll} Sxyz & \to xz(yz) \\ Kxy & \to x \\ Ix & \to x \end{array} \right\}$$

This paper shows a purely syntactical proof for an extension of Chew's theorem. This extension states that a weakly compatible TRS is **UN**. Roughly speaking, this weakly compatible condition allows a pair of nonroot overlapping rules if they return the same results (whereas a compatible condition allows only root overlapping rules). The main technique of this syntactical proof is an equational proof normalization called E-normalization that shows **UN** directly (not by way of the **CR** of a linearized TRS). This normalization technique differs those described in [2] and [9] in that it can be used even for TRSs that are nonterminating.



$$\theta(\beta) \;\leftarrow\; \theta(\alpha) \;=\; \sigma(\alpha) \;\to\; \sigma(\beta) \quad \Longrightarrow \quad \theta(\beta) \;=\; \sigma(\beta)$$

Fig. 1. Elimination of a reduction peak modulo equality on substitutions (Basic E-normalization)

Intuitively speaking, E-normalization is an elimination of reduction peaks modulo equality on substitutions (strongly overlapping cases, Fig. 1) and their variations (weakly compatible cases). Section 2 reviews the basic notation and terminology used in this paper, and it reviews results related to **CR**. Section 3 introduces a weight for proof structures that guarantees the termination of the E-normalization procedure. This weight is principally an extension of the parallel steps of an objective proof. Section 4 introduces E-normalization rules and their properties. Subsection 5.1 shows that any equality in a weakly compatible TRS has a proof in which there is no reduction peaks modulo equality (not restricted on substitutions. That is, two redexes are combined with equality that does not touch on their root symbols). Lemma 31 in subsection 5.2 then shows that a weakly compatible TRS is **UN**.

2 Term rewriting systems

Assuming that the reader is familiar with the basic TRSs concepts in [1] and [6], we briefly explain notations and definitions.

A term set $T(F, V)$ is a set of terms where F is a set of function symbols and V is a set of variable symbols. 0-ary function symbols are also called *constants*. The term set $T(F, V)$ may be abbreviated by simply T. A substitution θ is a map from V to $T(F, V)$. To avoid confusion with equality, we will use $\equiv$ to denote the syntactical identity between terms.

Definition 1. A position *position*(M, N) of a subterm N in a term M is defined by

$$\mathit{position}(M, N) = \left\{ \begin{array}{ll} \epsilon & \text{if } M \equiv N. \\ i \cdot u & \text{if } u = \mathit{position}(N_i, N) \quad \text{and} \quad M \equiv f(N_1, \cdots, N_n). \end{array} \right.$$

Let u and v be positions. We denote $u \prec v$ if $\exists w \neq \epsilon$ s.t. $v = u \cdot w$, we denote $u \preceq v$ if either $u = v$ or $u \prec v$, $u \parallel v$ if neither $u \preceq v$ nor $u \succeq v$, and we denote $u \not\parallel v$ if either $u \preceq v$ or $u \succeq v$.

For a set U of positions and a position v, we will denote $U \parallel v$ if $u \parallel v$ for $\forall u \in U$, we denote $U \not\parallel v$ if not $U \parallel v$, and we denote $v \prec U$ if $v \prec u$ for $\forall u \in U$.

The subterm N of M at position u is noted by M/u (i.e., $u = position(M, N)$). We say that u is a nonvariable position if M/u is not a variable. We denote a set of all positions in M by $pos(M)$, a set of all non-variable positions in M by $pos_F(M)$, and a set of positions of variables in M by $pos_V(M)$. A set of variables in M is denoted as $Var(M)$. A variable x is *linear* (in M) if x appears at most once in M. A variable x is *nonlinear* if it is not linear. A *replacement* of T/u with T', where u is a position in T, is denoted by $T[u \leftarrow T']$.

Definition 2. A finite set $R = \{\alpha_i \to \beta_i\}$ of ordered pairs of terms is said to be a *term rewriting system* (TRS) if each α_i is not a variable and $Var(\beta_i) \subseteq Var(\alpha_i)$. A binary relation called reduction is defined to be $M \to N$ if there exist a position u and a substitution θ s.t. $M/u \equiv \theta(\alpha_i)$ and $N \equiv M[u \leftarrow \theta(\beta_i)]$. A subterm $M/u \equiv \theta(\alpha_i)$ in M is said to be a *redex*. A *normal form* is a term that contains no redex. A set of normal forms of R is noted as $NF(R)$.

The symmetric closure of $\to$ is noted as $\leftrightarrow$. If a reduction $M \to N$ or $M \leftrightarrow N$ occurs at a position u, we will note $M \underset{u}{\to} N$ or $M \underset{u}{\leftrightarrow} N$. The reflexive transitive closure of $\to$ is noted as $\overset{*}{\to}$. An equality noted $=$ is the reflexive symmetric transitive closure of $\to$. A structure with an equality $=$ is said to be an equational system associated to a TRS R. We will use the default notations R for a TRS and E for an associated equational system.

Definition 3. A TRS R is *Church-Rosser* (**CR**) if $M = N$ implies $M \downarrow N$ (i.e., $\exists P$ s.t. $M \overset{*}{\to} P$ and $N \overset{*}{\to} P$). A TRS R is said to be *uniquely normalizing* (**UN**) if each set of equal terms has at most one normal form (i.e., $M = N$ and $M, N \in NF(R)$ imply $M \equiv N$).

Definition 4. A reduction rule is said to be *left linear* if any variable in its lhs is linear. A TRS is said to be *left linear* if all its reduction rules are left linear. We say a TRS is *nonlinear* if it is not left linear.

Definition 5. A pair of reduction rules $\alpha_i \to \beta_i$ and $\alpha_j \to \beta_j$ is said to be *overlapping* if there exist both a nonvariable position u in α_i and substitutions θ, σ s.t. $\theta(\alpha_i)/u \equiv \sigma(\alpha_j)$ (i.e., α_i/u and α_j are unifiable). We also say $\sigma(\alpha_j)$ overlaps with $\theta(\alpha_i)$ at u. A TRS R is said to be *nonoverlapping* if no pair of rules in R are overlapping except for trivial cases (i.e., $i = j \wedge u = \epsilon$), and a TRS R is said to be *strongly nonoverlapping* if R is nonoverlapping after renaming its nonlinear variables with fresh individual variables. A TRS R is said to be *weakly nonoverlapping* if any overlapping pair of rules returns same result (i.e., $\theta(\beta_i) \equiv \theta(\alpha_i)[u \leftarrow \sigma(\beta_j)]$).

Theorem 6 (Rosen[13]). *A left-linear weakly nonoverlapping TRS is* **CR**.

Corollary 7 ([11]). *A left-linear nonoverlapping TRS is* **CR**.

This theorem and its corollary intuitively rely on the commutativity of reductions, but nonlinear TRSs have more complex situations. Firstly, the commutativity of reductions is lost because it requires *synchronous* applications of a reduction on each occurrences of a variable (such as those applied in *term graph rewriting*[5]). Moreover, consider a variable x on the lhs of a rule: which x on the lhs is inherited to an occurrence of x's on the rhs? The following compatible conditions guarantee that if a pair of linearizations of rules is overlapping it returns same result under suitable combinations of variable-inheritances. Note that when a TRS is left-linear, a strongly nonoverlapping condition is the same as a nonoverlapping condition, and a weakly compatible condition is the same as a weakly nonoverlapping condition.

Definition 8 ([15]). (i) Let $\alpha \to \beta$ be a rewriting rule. A *linearization* of a term M is a term M' in which all the nonlinear variables in M are renamed to fresh individual variables. A *cluster* of a rewrite rule $\alpha \to \beta$ is $\{\alpha' \to \beta'_1, \cdots, \alpha' \to \beta'_n\}$, where α' is a linearization of α and each β'_i is a term obtained from β by replacing each nonlinear variable in α with one of the corresponding variables.

(ii) Let $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ be two rules of a TRS R. Let $\{\alpha'_1 \to \beta'_{11}, \cdots, \alpha'_1 \to \beta'_{1n}\}$ and $\{\alpha'_2 \to \beta'_{21}, \cdots, \alpha'_2 \to \beta'_{2m}\}$ be the two clusters corresponding to $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$. We say that $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ are *weakly compatible* if the following holds:

> *If $\sigma(\alpha'_2)$ overlaps with $\theta(\alpha'_1)$ at a position u, then the two clusters have a common instance wrt a context $C[\,] \equiv \alpha'_1[u \leftarrow \Box]$. That is,*
> $\{(\theta(\alpha'_1) \to \theta(\beta'_{1i})) \mid i = 1, \cdots, n\} \cap \{(\sigma(C[\alpha'_2]) \to \sigma(C[\beta'_{2j}])) \mid j = 1, \cdots, m\} \neq \phi.$

A TRS R is *weakly compatible* if all pairs of its reduction rules are weakly compatible.

(iii) Let notations be the same as in (ii), and let $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ be reduction rules. We say that $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ are *compatible* if they are weakly compatible and α'_1 and α'_2 may overlap only at the root (i.e., $u = \epsilon$). A TRS R is *compatible* if all pairs of reduction rules are compatible.

Example 1. Regarding a product xy in **CL-pc** as $apply(x, y)$, **CL-pc** is a compatible TRS. **CL-sp** (**CL** with surjective pairing[12, 15]) is not weakly compatible.

Theorem 9 (Chew[7]). *A compatible TRS is* **UN**.

Corollary 10 (Klop[3]). *A strongly nonoverlapping TRS is* **UN**.

3 Equational proof: structure and weight

3.1 Proof structure

Definition 11. A sequence of terms combined by the $\leftrightarrow$ relation is said to be a *proof structure*. A *proof* is a pair consisting of an equality and its proof structure, and it is denoted by $M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} \cdots \underset{u_{n-1}}{\leftrightarrow} M_{n-1} \underset{u_n}{\leftrightarrow} M_n \;\Rightarrow\; M_0 = M_n$, where a reduction $M_{i-1} \underset{u_i}{\leftrightarrow} M_i$ occurs at a position u_i. We will omit positions u_i if they are clear or unspecified from the context.

Instead of writing whole terms, we will use proof structure variables $\mathcal{P}, \mathcal{P}', \mathcal{P}_1, \mathcal{P}_2, \cdots$ for proof structures. For proofs $\mathcal{P}_1 \Rightarrow M_0 = M_1$, $\mathcal{P}_2 \Rightarrow M_1 = M_2$, $\cdots$, $\mathcal{P}_n \Rightarrow M_{n-1} = M_n$, a concatenation $\mathcal{P}_1, \mathcal{P}_2, \cdots, \mathcal{P}_n$ is denoted by $conc(\mathcal{P}_1, \cdots, \mathcal{P}_n)$, and the resulting proof is denoted by $conc(\mathcal{P}_1, \cdots, \mathcal{P}_n) \Rightarrow M_0 = M_n$. For a proof structure $\mathcal{P}$ of $M_1 \leftrightarrow M_2 \leftrightarrow \cdots \leftrightarrow M_n$ and a context $C[\,]$, we will denote $C[M_1] \leftrightarrow C[M_2] \leftrightarrow \cdots \leftrightarrow C[M_n]$ by $C[\mathcal{P}]$.

To emphasize the specific structure of a proof structure, we will also introduce two abbreviated notations: a sequence and a collection. A sequence of proofs $[\mathcal{P}_1 \Rightarrow M_0 = M_1\,;\, \cdots\,;\, \mathcal{P}_n \Rightarrow M_{n-1} = M_n]$ is a proof structure for $M_0 = M_n$. It emphasizes the specific intermediate terms $M_0, \cdots, M_n$. Its proper form is obtained by the unfolding rule (S):

> (S) *A sequence $[\mathcal{P}_1 \Rightarrow M_0 = M_1;\, \cdots\,;\, \mathcal{P}_n \Rightarrow M_{n-1} = M_n]$ is a proof structure for $M_0 = M_n$. It is unfolded to $conc(\mathcal{P}_1, \cdots, \mathcal{P}_n) \Rightarrow M_0 = N_n$.*

A collection of proofs $[\![\mathcal{P}_1 \Rightarrow M_1 = M'_1\,,\, \cdots\,,\, \mathcal{P}_n \Rightarrow M_n = M'_n]\!]$ is a proof structure for $C[M_1, \cdots, M_n] = C[M'_1, \cdots, M'_n]$. Consider an example of a proof structure $f(g(a), h(y)) \underset{1}{\to} f(h(a), h(y)) \underset{1}{\leftarrow} f(h(a), g(y))$ for a TRS $\{g(x) \to h(x)\}$. Parallel subproofs $g(a) \to h(a)$ and $h(y) \leftarrow g(y)$ induce an equality $f(g(a), h(y)) = f(h(a), g(y))$. A collection of proofs emphasizes such a situation. Its proper form is obtained by the unfolding rule (C):

> (C) *Let u_i be disjoint positions in $pos(M) \cap pos(N)$. A collection $[\![\mathcal{P}_1 \Rightarrow M/u_1 = N/u_1\,, \cdots\,,\, \mathcal{P}_n \Rightarrow M/u_n = N/u_n]\!]$ is a proof structure for $M = N$. It is unfolded to $conc(C_1[\mathcal{P}_1], \cdots, C_n[\mathcal{P}_n]) \Rightarrow M = N$, where $C_i[\Box] \equiv M[u_j \leftarrow N/u_j$ for $\forall j < i;\; u_i \leftarrow \Box]$.*

Our technique for transforming an equational proof to a *simpler* one is E-normalization. This normalization (introduced in Section 4) eliminates reduction peaks modulo equality on substitutions and their variations. For instance, consider a TRS $\{f(x,y) \rightarrow g(x,y,y), \cdots\}$. Assume there exist an equality $g(\theta(x),\theta(y),\theta(y)) = g(\sigma(x),\sigma(y),\sigma(y))$ that is proved by $g(\theta(x),\theta(y),\theta(y)) \leftarrow f(\theta(x),\theta(y))$, a collection of proofs $[\![\mathcal{P}_x \Rightarrow \theta(x)=\sigma(x), \mathcal{P}_y \Rightarrow \theta(y)=\sigma(y)]\!] \Rightarrow f(\theta(x),\theta(y)) = f(\sigma(x),\sigma(y))$, and $f(\sigma(x),\sigma(y)) \rightarrow g(\sigma(x),\sigma(y),\sigma(y))$. Then a (basic) E-normalization (Fig. 1) transforms it to a *simpler* proof $[\![\mathcal{P}_x \Rightarrow \theta(x)=\sigma(x), \mathcal{P}_y \Rightarrow \theta(y)=\sigma(y), \mathcal{P}_y \Rightarrow \theta(y)=\sigma(y)]\!] \Rightarrow g(\theta(x),\theta(y),\theta(y)) = g(\sigma(x),\sigma(y),\sigma(y))$.

To show the termination of E-normalization, the proof weight will be introduced in Subsection 3.2. This normalization decreases a parallel step of a transformed part in a proof structure, but when that part may be properly included in a whole proof, the number of parallel steps of a whole proof may not decrease. We therefore need an extension that is more sensitive to a parallel step of its proper substructure: a proof weight. Lemma 19 in Subsection 3.2 will show the termination of E-normalization procedure (Theorems 21 and 25).

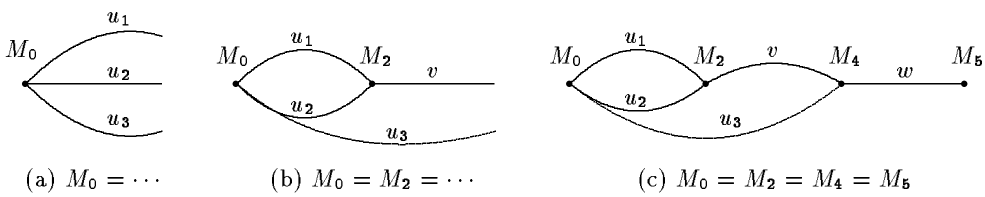


Fig. 2. Construction of E-graph for $M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} M_2 \underset{u_3}{\leftrightarrow} M_3 \underset{v}{\leftrightarrow} M_4 \underset{w}{\leftrightarrow} M_5$.

The graphical intuition of a proof weight is sketched by a notion of an E-graph[17]. Let $M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} M_2 \underset{u_3}{\leftrightarrow} M_3 \underset{v}{\leftrightarrow} M_4 \underset{w}{\leftrightarrow} M_5 \Rightarrow M_0 = M_5$. Assume that u_1, u_2, and u_3 are disjoint, and that $v \prec u_1, u_2$, $v \parallel u_3$, and $w \prec u_3, v$. Then, the E-graph is generated as follows. Since u_1, u_2, and u_3 are disjoint, the edges corresponding to u_1, u_2, and u_3 stem in parallel from a vertex M_0 (Fig. 2(a)). Since $v \prec u_1, u_2$, the edges labeled with u_1 and u_2 are collected to a vertex M_2 and an edge labeled v stems from M_2 (Fig. 2(b)). Finally, since $w \prec u_3, v$ and $u_3 \parallel v$, the edges labeled u_3 and v are collected to a vertex M_4 and an edge labeled w stems from M_4 to a vertex M_5 (Fig. 2(c)). A set of proof strings (Subsection 3.2) is a set of all cycle-free paths from M_0 to M_5, such as (u_1, v, w), (u_2, v, w), and (u_3, w). The proof weight of $\mathcal{P}$ is a multiset of their lengths $\{3, 3, 2\}$.

3.2 Weight for proofs

Definition 12. Let $(u_1, u_2, \cdots, u_n)$ be a sequence of positions and v be a position. An operation $(u_1, u_2, \cdots, u_n) \odot v$ returns a set defined by

$$\begin{cases} \{(u_1, u_2, \cdots, u_n, v)\} & \text{if } v \not\parallel u_n. \\ \{(u_1, u_2, \cdots, u_n), (u_1, u_2, \cdots, u_i, v)\} & \text{if } v \parallel u_j \text{ for } \forall j > i \text{ and } v \not\parallel u_i. \\ \{(u_1, u_2, \cdots, u_n)\} & \text{if } v \parallel u_i \text{ for } \forall i. \end{cases}$$

Definition 13. Let $\mathcal{S}$ be a set of sequences of positions, and v be a position. An operation $\mathcal{S} \oplus u$ is defined to be

$$\begin{cases} \bigcup_{\bar{u} \in \mathcal{S}} \bar{u} \odot v & \text{if } \exists \bar{u} \in \mathcal{S} \text{ s.t. } \bar{u} \not\parallel v, \text{ and} \\ \mathcal{S} \cup \{(v)\} & \text{otherwise.} \end{cases}$$

Here $\bar{u} \parallel v$ if $u_i \parallel v$ for $\forall u_i \in \bar{u} = (u_1, u_2, \cdots, u_n)$, and $\bar{u} \not\parallel v$ if not $\bar{u} \parallel v$.

Definition 14. Let $M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} \cdots \underset{u_{n-1}}{\leftrightarrow} M_{n-1} \underset{u_n}{\leftrightarrow} M_n \Rightarrow M_0 = M_n$ be a proof. A *proof string* is an element of a set $\mathcal{S}(M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} \cdots \underset{u_{n-1}}{\leftrightarrow} M_{n-1} \underset{u_n}{\leftrightarrow} M_n)$ inductively defined to be

$$\begin{cases} \phi & \text{if } n = 0, \text{ and} \\ \mathcal{S}(M_0 \underset{u_1}{\leftrightarrow} M_1 \underset{u_2}{\leftrightarrow} \cdots \underset{u_{n-1}}{\leftrightarrow} M_{n-1}) \oplus u_n & \text{otherwise.} \end{cases}$$

For two sets $\mathcal{S}_1$, $\mathcal{S}_2$ of sequences of positions, $\mathcal{S}_1 \otimes \mathcal{S}_2$ is defined to be $\cup_{(v_1,\cdots,v_n)\in\mathcal{S}_2} \mathcal{S}_1 \oplus v_1 \oplus \cdots \oplus v_n$. Thus $\mathcal{S}(conc(\mathcal{P}_1, \mathcal{P}_2)) = \mathcal{S}(\mathcal{P}_1) \otimes \mathcal{S}(\mathcal{P}_2)$.

Definition 15. Let $\mathcal{P} \Rightarrow M = N$ be a proof. A *proof weight* $w(\mathcal{S}(\mathcal{P}))$ is a multiset of string-lengths of proof strings in $\mathcal{S}(\mathcal{P})$ where a string-length $l(\bar{u})$ is a number n for $\bar{u} = (u_1, u_2, \cdots, u_n)$. The ordering on proof weights is defined to be the multiset extension of the usual ordering on natural numbers[10].

Definition 16. Let $\mathcal{P} \Rightarrow M = N$ be a proof. A *boundary* $\partial\mathcal{P}$ is a set of positions defined by $min(\{v \mid v \text{ appears in } \bar{u} \in \mathcal{S}(\mathcal{P})\})$ (i.e., a set of minimum positions appeared in each proof string in $\mathcal{S}(\mathcal{P})$). If a proof structure $\mathcal{P}$ satisfies $\forall v \in \partial\mathcal{P}$ s.t. $w \prec v$, we denote $\mathcal{P}$ by $M \underset{w\prec}{=} N$. If $\forall v \in \partial\mathcal{P}$ s.t. $w \preceq v$, we denote $\mathcal{P}$ by $M \underset{w\preceq}{=} N$.

Definition 17. Let $\mathcal{S}$ be a set of proof strings. $\mathcal{S}^u$ is defined to be $\{\bar{v} \in \mathcal{S} \mid \bar{v} \not\parallel u\}$ for a position u. For a set of disjoint positions U, $\mathcal{S}^U$ is $\cup_{u\in U}\mathcal{S}^u$. For positions v, w s.t. $v \preceq w$, $\mathcal{S}^{v/w}$ is $\mathcal{S}^v - \mathcal{S}^w$. For a position v and a set of disjoint positions W, $\mathcal{S}^{v/W}$ is $\mathcal{S}^v - \mathcal{S}^W$.

Definition 18. Let $\bar{w} = (w_1, \cdots, w_m)$ be a proof string and u be a position. We will denote $(w_1, \cdots, w_i)$ by $\bar{w} \lhd u$ if $w_i \not\parallel u$ and $w_j \parallel u$ for $\forall j > i$, and we will denote $(w_i, \cdots, w_n)$ by $u \rhd \bar{w}$ if $w_i \not\parallel u$ and $w_j \parallel u$ for $\forall j < i$. For a set $\mathcal{S}$ of proof strings, $\mathcal{S} \lhd u$ is $\{\bar{w} \lhd u \mid \bar{w} \in \mathcal{S}\}$ and $u \rhd \mathcal{S}$ is $\{u \rhd \bar{w} \mid \bar{w} \in \mathcal{S}\}$.

Lemma 19. Let an equality $S = T$ have two proof structures $\mathcal{P}_1, \mathcal{P}_2$, and let u, v be positions s.t. $u \not\parallel v$. Assume $\mathcal{P}_1$ has a form $S \underset{u}{\leftrightarrow} S' \underset{max(u,v)\prec}{=} T' \underset{v}{\leftrightarrow} T$ and $\mathcal{P}_2$ has a form $S \underset{max(u,v)\prec}{=} T$. Then $max(w(\mathcal{S}(\mathcal{P}_1))) > max(w(\mathcal{S}(\mathcal{P}_2)))$ implies $w(\mathcal{S}(\mathcal{P}_1^*)) \gg w(\mathcal{S}(\mathcal{P}_2^*))$, where

(i) $\mathcal{P}_1^*$ is $[\mathcal{P} \Rightarrow M = S;\ \mathcal{P}_1 \Rightarrow S = T;\ \mathcal{P}' \Rightarrow T = N\ ]$,
(ii) $\mathcal{P}_2^*$ is $[\mathcal{P} \Rightarrow M = S;\ \mathcal{P}_2 \Rightarrow S = T;\ \mathcal{P}' \Rightarrow T = N\ ]$, and
(iii) $\gg$ is the multiset extension of the ordering $>$ on natural numbers.

Proof Without loss of generality, we can assume $u \preceq v$ and $max(u, v) = v$. We remark that if positions u, v satisfy $u \preceq v$, then, for any set $\mathcal{S}$ of proof strings, $\mathcal{S}^u \supseteq \mathcal{S}^v$, $max(w(\mathcal{S} \lhd u)) \geq max(w(\mathcal{S} \lhd v))$, and $max(w(u \rhd \mathcal{S})) \geq max(w(v \rhd \mathcal{S}))$.

We will estimate the maximum length in $w(\mathcal{S}(\mathcal{P}_1^*) - \mathcal{S}(\mathcal{P}_1^*) \cap \mathcal{S}(\mathcal{P}_2^*))$ and $w(\mathcal{S}(\mathcal{P}_2^*) - \mathcal{S}(\mathcal{P}_1^*) \cap \mathcal{S}(\mathcal{P}_2^*))$. The former will be shown to be greater than the latter, and this will show that $w(\mathcal{S}(\mathcal{P}_1^*) \gg w(\mathcal{S}(\mathcal{P}_2^*)$.

(1) A proof string in $\mathcal{S}(\mathcal{P}_1^*) = \mathcal{S}(\mathcal{P}) \otimes \mathcal{S}(\mathcal{P}_1) \otimes \mathcal{S}(\mathcal{P}')$ has one of following forms:

(1-i)	$(w_1, \cdots, w_i, u, u_1, \cdots, u_s, v, w'_j, \cdots, w'_n)$	for	$(w_1, \cdots, w_i) \in \mathcal{S}(\mathcal{P})^u \lhd u$
			$(u, u_1, \cdots, u_s, v) \in \mathcal{S}(\mathcal{P}_1)$
			$(w'_j, \cdots, w'_n) \in v \rhd \mathcal{S}(\mathcal{P}')^v$
(1-ii)	$(w_1, \cdots, w_i, u, w'_j, \cdots, w'_n)$	for	$(w_1, \cdots, w_i) \in \mathcal{S}(\mathcal{P})^u \lhd u$
			$(w'_j, \cdots, w'_n) \in u \rhd \mathcal{S}(\mathcal{P}')^{u/v}$
(1-iii)	$(w_1, \cdots, w_i, w'_j, \cdots, w'_n)$	for	$(w_1, \cdots, w_i) \in \mathcal{S}(\mathcal{P})^{\epsilon/u} \lhd w'_j$
			$(w'_j, \cdots, w'_n) \in w_i \rhd \mathcal{S}(\mathcal{P}')^{\epsilon/u}$

These sets are respectively denoted $\mathcal{S}_{(1-i)}$, $\mathcal{S}_{(1-ii)}$, and $\mathcal{S}_{(1-iii)}$. The maximum length of $\mathcal{S}_{(1-i)}$ and $\mathcal{S}_{(1-ii)}$ are estimated as

$max(w(\mathcal{S}_{(1-i)})) = max(w(\mathcal{S}(\mathcal{P}) \lhd u)) + max(w(\mathcal{S}(\mathcal{P}))) + max(w(v \rhd \mathcal{S}(\mathcal{P}')))$ and
$max(w(\mathcal{S}_{(1-ii)})) = max(w(\mathcal{S}(\mathcal{P}) \lhd u)) + 1 + max(w(u \rhd \mathcal{S}(\mathcal{P}')^{u/v}))$.

(2) A proof string in $\mathcal{S}(\mathcal{P}_2^*) = \mathcal{S}(\mathcal{P}) \otimes \mathcal{S}(\mathcal{P}_2) \otimes \mathcal{S}(\mathcal{P}')$ has one of following forms:

$$
\begin{array}{lll}
\text{(2-i)} & (w_1,\cdots,w_i,u'_1,\cdots,u'_s,w'_j,\cdots,w'_n) & \text{for } (w_1,\cdots,w_i)\in\mathcal{S}(\mathcal{P})^{u'_1}\lhd u'_1 \\
& & \quad (u'_1,\cdots,u'_s)\in w_i\ \rhd\mathcal{S}(\mathcal{P}_2)\lhd w'_j \\
& & \quad (w'_j,\cdots,w'_n)\in u'_s\ \rhd\mathcal{S}(\mathcal{P}')^{u'_s} \\
\text{(2-ii)} & (w_1,\cdots,w_i,w'_j,\cdots,w'_n) & \text{for } (w_1,\cdots,w_i)\in\mathcal{S}(\mathcal{P})^{\epsilon/\partial\mathcal{P}_2}\lhd w'_j \\
& & \quad (w'_j,\cdots,w'_n)\in w_i\ \rhd\mathcal{S}(\mathcal{P}')^{\epsilon/\partial\mathcal{P}_2}
\end{array}
$$

These sets are denoted $\mathcal{S}_{(2-i)}$ and $\mathcal{S}_{(2-ii)}$. Then $\mathcal{S}(\mathcal{P}_1^*)\cap\mathcal{S}(\mathcal{P}_2^*)=\mathcal{S}_{(1-iii)}$ and $\mathcal{S}_{(2-i)}\cap\mathcal{S}_{(1-iii)}=\phi$. A set $\mathcal{S}_{(2-ii)'}=\mathcal{S}_{(2-ii)}-\mathcal{S}_{(1-iii)}$ consists of elements of the form

$$
\begin{array}{lll}
\text{(2-ii)'} & (w_1,\cdots,w_i,w'_j,\cdots,w'_n) & \text{for } (w_1,\cdots,w_i)\in\mathcal{S}(\mathcal{P})^{u/\partial\mathcal{P}_2}\lhd w'_j \\
& & \quad (w'_j,\cdots,w'_n)\in w_i\ \rhd\mathcal{S}(\mathcal{P}')^{u/\partial\mathcal{P}_2}.
\end{array}
$$

Since $u\rhd\mathcal{S}(\mathcal{P}')^{u/\partial\mathcal{P}_2}=u\rhd\mathcal{S}(\mathcal{P}')^{u/v}\cup u\rhd\mathcal{S}(\mathcal{P}')^{v/\partial\mathcal{P}_2}$, the maximum lengths $max(w(\mathcal{S}_{(2-i)}))$ and $max(w(\mathcal{S}_{(2-ii)'}))$ are estimated as follows:

$$
\begin{array}{rl}
max(w(\mathcal{S}_{(2-i)})) & \le max(\ w(\mathcal{S}(\mathcal{P})^{\partial\mathcal{P}_2}\lhd u))+max(w(\mathcal{S}(\mathcal{P}_2)))+max(w(v\ \rhd\mathcal{S}(\mathcal{P}')^{\partial\mathcal{P}_2})) \\
& < max(\ w(\mathcal{S}(\mathcal{P})\lhd u))+max(w(\mathcal{S}(\mathcal{P}_1)))+max(w(v\ \rhd\mathcal{S}(\mathcal{P}'))) \\
& = max(\ w(\mathcal{S}_{(1-i)})) \\
max(w(\mathcal{S}_{(2-ii)'})) & \le max(\ w(\mathcal{S}(\mathcal{P})^{u/\partial\mathcal{P}_2}\lhd u)+max(w(u\ \rhd\mathcal{S}(\mathcal{P}')^{u/\partial\mathcal{P}_2})) \\
& \le max(\ max(w(\mathcal{S}(\mathcal{P})^{u/v}\lhd u)+max(w(u\ \rhd\mathcal{S}(\mathcal{P}')^{u/v})), \\
& \qquad max(w(\mathcal{S}(\mathcal{P})^{v/\partial\mathcal{P}_2}\lhd u)+max(w(u\ \rhd\mathcal{S}(\mathcal{P}')^{v/\partial\mathcal{P}_2}))\ ) \\
& < max(\ max(w(\mathcal{S}_{(1-i)})),max(w(\mathcal{S}_{(1-ii)}))\ )
\end{array}
$$

Thus $w(\mathcal{S}(\mathcal{P}_1^*))\gg w(\mathcal{S}(\mathcal{P}_2^*))$. ■

4 *E*-normalization

4.1 Basic *E*-normalization

Definition 20. Let $\alpha\to\beta$ be a reduction rule, and let θ and σ be substitutions. Then, the following transformation rule (from the upper column to the lower column) is said to be a *basic E-normalization* rule:

$$
\frac{\left[\begin{array}{l}\theta(\beta)\leftarrow\theta(\alpha)\ \Rightarrow\ \theta(\beta)=\theta(\alpha)\ ; \\ [\![\mathcal{P}_u\ \Rightarrow\ \theta(\alpha)/u=\sigma(\alpha)/u \text{ for } \forall u\in pos_V(\alpha)]\!]\ \Rightarrow\ \theta(\alpha)=\sigma(\alpha)\ ; \\ \sigma(\alpha)\to\sigma(\beta)\ \Rightarrow\ \sigma(\alpha)=\sigma(\beta)\end{array}\right]\Rightarrow\theta(\beta)=\sigma(\beta)}{[\![\mathcal{P}_v\ \Rightarrow\ \theta(\beta)/v=\sigma(\beta)/v \text{ for } \forall v\in pos_V(\beta)]\!]\qquad\Rightarrow\theta(\beta)=\sigma(\beta)}
$$

if a proof structure $\mathcal{P}_v$ in the lower column is arbitrarily selected from a set of $\mathcal{P}_u$'s in the upper column s.t. $\alpha/u\equiv\beta/v\in Var(\beta)$.

Theorem 21. *Each step of a basic E-normalization decreases the proof weight. Thus a basic E-normalization procedure for any proof always terminates.*

4.2 *E*-normalization for a weakly compatible TRS

Definition 22. Let $\alpha_1\to\beta_1$ and $\alpha_2\to\beta_2$ be reduction rules, let θ and σ be substitutions, and let $C[\,]\equiv\theta(\alpha_1)[u\leftarrow\Box]$. Assume that $\theta(\alpha_1)$ and $C[\sigma(\alpha_2)]$ are combined by a proof structure $\theta(\alpha)\underset{u\preceq}{=}C[\sigma(\alpha_2)]$. If $\partial(\theta(\alpha_1)\underset{u\prec}{=}C[\sigma(\alpha_2)])\cap pos_F(\alpha_1)\cap u\cdot pos_F(\alpha_2)=\phi$, then $\theta(\alpha_1)$ and $C[\sigma(\alpha_2)]$ are said to be *quasi-E-normalizable*.

Assume S and S' to be quasi-E-normalizable. The next lemma shows that if a TRS is weakly compatible, then any pair of reduction $T\leftarrow S$ and $S'\to T'$ in a proof can be eliminated by an E-normalization.

Lemma 23. Let a TRS R be weakly compatible. Assume that $\alpha_1\to\beta_1$ and $\alpha_2\to\beta_2$ are rules in R, that θ and σ be substitutions, and that a position $u\in pos_F(\alpha_1)$ s.t.

$[\![\mathcal{P}_w \Rightarrow \theta(\alpha_1)/u\cdot w = \sigma(\alpha_2)/w \text{ for } \forall w \in min(pos_V(\alpha_1/u)\cup pos_V(\alpha_2))]\!] \Rightarrow\ \theta(\alpha_1) = \sigma(C[\alpha_2])$ for $C[\,] \equiv \alpha_1[u \leftarrow \Box]$. Then linearizations of α_1/u and α_2 overlap, and $[\![\mathcal{P}_{w'} \Rightarrow \theta(\beta_1)/u\cdot w' = \sigma(\beta_2)/w' \text{ for } \forall w' \in min(pos_V(\beta_1/u)\cup pos_V(\beta_2))]\!] \Rightarrow \theta(\beta_1) = \sigma(C[\beta_2])$

Proof Let $\{\alpha_1' \to \beta_{11}', \cdots, \alpha_1' \to \beta_{1n}'\}$ and $\{\alpha_2' \to \beta_{21}', \cdots, \alpha_2' \to \beta_{2m}'\}$ be clusters of $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$. Let $W = min(pos_V(\alpha_1/u)\cup pos_V(\alpha_2))$, and let $C'[\,]$ be a linearization of $C[\,]$. Set substitutions θ' for α_1' and σ' for $C'[\alpha_2']$ as follows:

For a variable x in α_1' s.t. $w = position(\alpha_1', x)$,

$$\theta'(x) \equiv \begin{cases} \sigma(C[\alpha_2]/w) & \text{if } w \in u\cdot(W \cap pos_F(\alpha_2)), \quad \text{and} \\ \theta(\alpha_1/w) & \text{otherwise.} \end{cases}$$

For a variable y in $C'[\alpha_2']$ s.t. $w = position(C'[\alpha_2'], y)$,

$$\sigma'(y) \equiv \begin{cases} \theta(\alpha_1/w) & \text{if } w \in u\cdot(W - pos_F(\alpha_2)), \text{and} \\ \sigma(C[\alpha_2]/w) & \text{otherwise.} \end{cases}$$

Then linearizations α_1' and α_2' overlap at a position u (i.e., $\theta'(\alpha_1') \equiv \sigma'(C[\alpha_2'])$). Since R is weakly compatible, there exists the intersection $(\theta'(\alpha_1') \to \theta'(\beta_{1s}')) \equiv (\sigma'(C[\alpha_2']) \to \sigma'(C[\beta_{2t}']))$ between two clusters (i.e., $\theta'(\beta_{1s}') \equiv \sigma'(C[\beta_{2t}'])$). Thus a set of proofs $\theta(\alpha_1/u\cdot w) = \sigma(\alpha_2/w)$ for $\forall w \in min(pos_V(\alpha_1/u)\cup pos_V(\alpha_2))$ also combines $\theta(\beta_1)$ and $\sigma(C[\beta_2])$ as $\theta(\beta_1) = \theta'(\beta_{1s}') \equiv \sigma'(C[\beta_{2t}']) = \sigma(C[\beta_2])$. ∎

Definition 24. Let $\alpha_1 \to \beta_1$ and $\alpha_2 \to \beta_2$ be weakly compatible. Assume α_1'/u and α_2' to be unified for $u \in pos_F(\alpha_1')$, where α_1' and α_2' are the linearizations of α_1 and α_2. Let $C[\,]$ be a context that has a hole $\Box$ at u s.t. $C[\alpha_2'] \equiv \alpha_1'$, and let θ and σ be substitutions. The following transformation rules are said to be the *E-normalization* rules:

$$\frac{\left[\begin{array}{l} \theta(\beta_1) \leftarrow \theta(\alpha_1) \;\Rightarrow\; \theta(\beta_1) = \theta(\alpha_1)\ ; \\ \left[\!\!\left[\begin{array}{l} \mathcal{P}_w \;\Rightarrow\; \theta(\alpha_1)/u\cdot w = \sigma(\alpha_2)/w \\ \text{for } \forall w \in min(pos_V(\alpha_1/u)\cup pos_V(\alpha_2)) \end{array}\right]\!\!\right] \Rightarrow\; \theta(\alpha_1) = \sigma(C[\alpha_2]); \\ \sigma(\alpha_2) \to \sigma(\beta_2) \;\Rightarrow\; \sigma(C[\alpha_2]) = \sigma(C[\beta_2]) \end{array}\right] \Rightarrow \theta(\beta_1) = \sigma(C[\beta_2])}{\left[\!\!\left[\begin{array}{l} \mathcal{P}_{w'} \;\Rightarrow\; \theta(\beta_1/u)\cdot w' = \sigma(\beta_2)/w' \\ \text{for } \forall w' \in min(pos_V(\beta_1/u)\cup pos_V(\beta_2)) \end{array}\right]\!\!\right] \Rightarrow \theta(\beta_1) = \sigma(C[\beta_2])}$$

and its inverted form

$$\frac{\left[\begin{array}{l} \sigma(\beta_2) \leftarrow \sigma(\alpha_2) \;\Rightarrow\; \sigma(C[\beta_2]) = \sigma(C[\alpha_2]) \\ \left[\!\!\left[\begin{array}{l} \mathcal{P}_w \;\Rightarrow\; \sigma(\alpha_2)/w = \theta(\alpha_1)/u\cdot w \\ \text{for } \forall w \in min(pos_V(\alpha_1/u)\cup pos_V(\alpha_2)) \end{array}\right]\!\!\right] \Rightarrow\; \sigma(C[\alpha_2]) = \theta(\alpha_1); \\ \theta(\alpha_1) \to \theta(\beta_1) \;\Rightarrow\; \theta(\alpha_1) = \theta(\beta_1) \end{array}\right] \Rightarrow \sigma(C[\beta_2]) = \theta(\beta_1)}{\left[\!\!\left[\begin{array}{l} \mathcal{P}_{w'} \;\Rightarrow\; \sigma(\beta_2)/w' = \theta(\beta_1/u)\cdot w' \\ \text{for } \forall w' \in min(pos_V(\beta_1/u)\cup pos_V(\beta_2)) \end{array}\right]\!\!\right] \Rightarrow \sigma(C[\beta_2]) = \theta(\beta_1)}$$

Here a proof structure $\mathcal{P}_{w'}$ in the lower column is arbitrarily selected from a set of $\mathcal{P}_w$ in the upper column s.t. $\theta(\alpha_1)/u\cdot w \equiv \theta(\beta_1)/u\cdot w'$ and $\sigma(\alpha_2)/w \equiv \sigma(\beta_2)/w'$.

Theorem 25. *Let R be a weakly compatible TRS. Each step of an E-normalization decreases the proof weight. Thus the E-normalization procedure for any proof always terminates.*

Example 2. For a product xy such as $apply(x, y)$, CL-pc is a compatible TRS. CL-pc has overlapping rules such as $(CTxy \to x, Czxx \to x)$ and $(CFxy \to x, Czxx \to x)$. The following rule is an example of an E-normalization rule in addition to the basic E-normalization rules:

$$\frac{\left[\begin{array}{ll} \theta(x) \leftarrow C\ T\ \theta(x)\ \theta(y) & \Rightarrow \theta(x) = C\ T\ \theta(x)\ \theta(y)\ ; \\ \left[\!\!\left[\begin{array}{l} \mathcal{P}_1 \Rightarrow\ T = \sigma(z'), \\ \mathcal{P}_2 \Rightarrow\ \theta(x) = \sigma(x'), \\ \mathcal{P}_3 \Rightarrow\ \theta(y) = \sigma(x') \end{array}\right]\!\!\right] & \Rightarrow C\ T\ \theta(x)\ \theta(y) = C\ \sigma(z')\ \sigma(x')\ \sigma(x')\ ; \\ C\ \sigma(z')\ \sigma(x')\ \sigma(x') \to \sigma(x') \Rightarrow C\ \sigma(z')\ \sigma(x')\ \sigma(x') = \sigma(x') & \end{array}\right] \Rightarrow \theta(x) = \sigma(x')}{\mathcal{P}_2\ \Rightarrow \theta(x) = \sigma(x')}$$

for substitutions θ and σ.

5 Sufficient condition for the UN property

5.1 E-overlapping pair

Definition 26. Let $\mathcal{P} \Rightarrow M_0 = M_n$ be a proof. Assume $\mathcal{P}$ has the form
$M_0 \leftrightarrow M_1 \leftrightarrow \cdots \leftrightarrow M_{i-1} \underset{u}{\leftarrow} M_i \underset{max(u,v)\prec}{=} M_j \underset{v}{\rightarrow} M_{j+1} \leftrightarrow \cdots \leftrightarrow M_{n-1} \leftrightarrow M_n$.
Then a pair of terms (M_i, M_j) is said to be an *E-overlapping pair*.

Definition 27. A proof is *E-normal* if no E-normalization rule is applicable. An equational proof is said to be *E-overlapping-pair-free* if there are no E-overlapping pairs in it. A TRS is *E-nonoverlapping* if every E-normal proof is E-overlapping-pair-free.

Theorem 28. *If a TRS R is weakly compatible, then R is E-nonoverlapping.*

Proof Let $\mathcal{P} \Rightarrow M = N$ be an E-normal proof. Assume there exists an E-overlapping pair in $\mathcal{P}$. Let (S,T) be an innermost E-overlapping pair in $\mathcal{P}$. Then (S,T) is quasi-E-normalizable (otherwise there is an inner E-overlapping pair between S and T), and E-normalization can be applicable to (S,T) from Lemma 23. This contradicts the E-normal assumption on $\mathcal{P}$. ■

5.2 The UN property for an E-nonoverlapping TRS

Definition 29 [11]. A term-length $\Delta(M)$ for a term M is defined to be

$$\begin{cases} \Delta(x) = 1 & \text{for a variable } x, \quad \text{and} \\ \Delta(f(M_1,\cdots,M_n)) = 1 + \sum_{i=1}^{n} \Delta(M_i) & \text{otherwise.} \end{cases}$$

Proposition 30. Assume that a TRS R is E-nonoverlapping. If $\mathcal{P} \Rightarrow M = N$ is an E-normal proof and $M \not\equiv N$, then there exist a position $u \in \partial\mathcal{P}$, a substitution σ, and a rule $\alpha \rightarrow \beta$ s.t. $\mathcal{P}$ is either $M = N[u \leftarrow \sigma(\beta)] \underset{u}{\leftarrow} N[u \leftarrow \sigma(\alpha)] \underset{u\prec}{=} N$, or $M \underset{u\prec}{=} M[u \leftarrow \sigma(\alpha)] \underset{u}{\rightarrow} M[u \leftarrow \sigma(\beta)] = N$.

Proof $\mathcal{P}$ is a nontrivial proof because $M \not\equiv N$. If $\mathcal{P}$ holds for neither of the cases specified in Proposition 30, there exist a position $u \in \mathcal{P}$, a substitution σ', and a rule $\alpha' \rightarrow \beta'$ s.t. $\mathcal{P}$ includes $M/u \underset{\epsilon\prec}{=} \sigma(\beta) \underset{\epsilon}{\leftarrow} \sigma(\alpha) = \sigma'(\alpha') \underset{\epsilon}{\rightarrow} \sigma'(\beta') \underset{u\prec}{=} N/u$. Then there must exist an E-overlapping pair in $\sigma(\beta) \underset{\epsilon}{\leftarrow} \sigma(\alpha) = \sigma'(\alpha') \underset{\epsilon}{\rightarrow} \sigma'(\beta')$. This is contradiction. ■

Lemma 31. Assume that a TRS R is weakly compatible. Then

(i) If $M \underset{\epsilon\prec}{=} \theta(\alpha) \rightarrow \theta(\beta)$ is an E-normal proof for some rule $\alpha \rightarrow \beta \in R$ and a substitution θ, then $M \notin NF(R)$.
(ii) If $N_1 = N_2$ and $N_1, N_2 \in NF(R)$, then $N_1 \equiv N_2$.

Proof Since R is weakly compatible, Theorem 28 implies that R is E-nonoverlapping. Assume that (i) of Lemma 31 holds, then Theorem 25 and Proposition 30 imply that if $N_1 = N_2$ and $N_1 \not\equiv N_2$, then either N_1 or N_2 is not a normal form. Thus if (i) holds for $\Delta(M) < n$, (ii) also holds for $\Delta(N_1), \Delta(N_2) < n$.

The proof of (i) is due to the induction on $n = \Delta(M)$. If $n = 1$, M must be a constant or a variable. Then $M \underset{\epsilon\prec}{=} \theta(\alpha)$ implies $M \equiv \theta(\alpha)$, and M is a redex. Let $n > 1$ and $M \underset{\epsilon\prec}{=} \theta(\alpha) \rightarrow \theta(\beta) \Rightarrow M = \theta(\beta)$ be an E-normal proof.

Assume $M \in NF(R)$. Thus $M \not\equiv \theta(\alpha)$ and according to Proposition 30 there are a position $u \in \partial(M \underset{\epsilon\prec}{=} \theta(\alpha))$, a substitution σ, and a rule $\alpha' \rightarrow \beta' \in R$ s.t. $M \underset{\epsilon\prec}{=} \theta(\alpha) \rightarrow \theta(\beta)$ has one of the following forms:

$$\begin{cases} (1)\ M \underset{u\prec}{=} M[u \leftarrow \sigma(\alpha')] \underset{u}{\rightarrow} M[u \leftarrow \sigma(\beta')] = \theta(\alpha) \rightarrow \theta(\beta) \\ (2)\ M = \theta(\alpha)[u \leftarrow \sigma(\beta')] \underset{u}{\leftarrow} \theta(\alpha)[u \leftarrow \sigma(\alpha')] \underset{u\prec}{=} \theta(\alpha) \rightarrow \theta(\beta) \end{cases}$$

Form (1) satisfies $\Delta(M/u) < \Delta(M)$, and $M/u \underset{\epsilon\prec}{=} \sigma(\alpha') \to \sigma(\beta')$ is E-normal (because a subproof of an E-normal proof is E-normal). Thus $M/u \notin NF(R)$ from the induction hypothesis. This implies $M \notin NF(R)$.

Let us divide Form (2) into two cases: (a) $u \in pos_F(\alpha)$ and (b) $u \notin pos_F(\alpha)$.

(a) Since $u \in pos_F(\alpha)$, $(\theta(\alpha)[u \leftarrow \theta(\alpha')], \theta(\alpha))$ is an E-overlapping pair. Every E-normal proof in R is E-overlapping-pair-free. Thus this is a contradiction.

(b) There exist a nonlinear variable x in α and positions $v_1, v_2 \in position(\alpha, x)$ s.t. $M/v_1 \not\equiv M/v_2$ and $M/v_1 = M/v_2$. (Otherwise, M is a redex for $\alpha \to \beta$.) If $M \in NF(R)$, then $M/v_1, M/v_2 \in NF(R)$ and applying the induction hypothesis to (ii) shows $M/v_1 \equiv M/v_2$. This is a contradiction. Thus $M \notin NF(R)$. ■

Theorem 32. *If a TRS R is weakly compatible, then R is* **UN**.

Corollary 33. *Let R be a weakly compatible TRS. If R has at least two distinct normal forms, then an associated equality system E is consistent.*

Acknowledgments

I thank Dr. Yoshihito Toyama (NTT Communication Science Lab.), Dr. Aart Middeldorp (Hitachi Advanced Research Lab.), and Professor Michio Oyamaguchi (Mie Univ.) for valuable comments and discussions. I also thank my supervisor Hirofumi Katsuno for his careful reading, Professor Touru Naoi (Gifu Univ.) for informing me of references, and Dr. Satoshi Ono (NTT Software Lab.) for suggestions.

References

1. J.van Leeuwen (eds.), "Handbook of Theoretical Computer Science, Vol.B Formal models and Semantics", The MIT Press/Elservier (1990)
2. L.Bachmair, "Canonical Equational Proofs," Birkhäuser (1991)
3. J.W.Klop, "Combinatory Reduction Systems", Mathematical Centre Tracts 127, CWI Amsterdam (1980)
4. D.S.Johnson (eds.), "STOC/FOCS Bibliography (Preliminary version)", ACM SIGACT (1991)
5. Barendregt,H.P., M.C.J.D.Eekelen, J.R.W.Glauert, J.R.Kennaway, M.J.Pasmeijer, and M.R.Sleep, "Term Graph Rewriting", Proc. Parallel Architectures and Languages Europe Vol.2, pp.141-158, Springer LNCS 259 (1987)
6. J.A.Bergstra,and J.W.Klop, "Conditional Rewrite Rules : Confluence and Termination," Journal of Computer and System Science, Vol.32, pp.323-362 (1986)
7. P.Chew, "Unique Normal Forms in Term Rewriting Systems with Repeated Variables",Proc. the 13th ACM Sympo. on Theory of Computing, pp.7-18 (1981)
8. P.L.Curien,and G.Ghelli, "On Confluence for Weakly Normalizing Systems," Proc. the 4th Int. Conf. on Reduction Techniques and Applications, pp.215-225, Springer LNCS 488 (1991)
9. N.Dershowitz,and M.Okada, "Proof-theoretic techniques for term rewriting theory," Proc. the 3rd IEEE Sympo. on Logic in Computer Science, pp.104-111 (1988)
10. N.Dershowitz, and Z.Manna, "Proving Termination with Multiset Orderings," Comm. of the ACM, Vol.22, No.8, pp.465-476 (1979)
11. G.Huet, "Confluent reductions : Abstract properties and applications to term rewriting systems", Journal of the ACM, Vol.27,No.4, pp.797-821 (1980)
12. J.W.Klop, and R.C.de Vrijer, "Unique normal forms for lambda calculus with surjective pairing," Information and Computation, Vol.80, pp.97-113 (1989)
13. B.K.Rosen, "Tree-manipulating systems and Church-Rosser theorems", Journal of the ACM, Vol.20, No.1, pp.160-187 (1973)
14. R.C.de Vrijer, "Extending the Lambda Calculus with Surjective Pairing is Conservative," Proc. the 4th IEEE Sympo. on Logic in Computer Science, pp.204-215 (1989)
15. R.C.de Vrijer, "Unique Normal Forms for Combinatory Logic with Parallel Conditional, a case study in conditional rewriting," Technical Report, Free University, Amsterdam (1990)
16. M.Ogawa,and S.Ono,"On the uniquely converging property of nonlinear term rewriting systems," IEICE Technical Report COMP89-7, pp.61-70 (1989)
17. M.Oyamaguchi, *Private Communications* (April 1992)