1991

A. Middeldorp, M. Starčević

A rewrite approach to polynomial ideal theory

Computer Science/Department of Software Technology Report CS-R9160 December

CWI, nationaal instituut voor onderzoek op het gebied van wiskunde en informatica

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

A Rewrite Approach to Polynomial Ideal Theory

Aart Middeldorp[1]

Department of Software Technology
CWI, Kruislaan 413, 1098 SJ Amsterdam
ami@cwi.nl

Mirjana Starčević[2]

Department of Mathematics and Computer Science
Vrije Universiteit, de Boelelaan 1081a, 1081 HV Amsterdam
mirjana@cs.vu.nl

December 19, 1991

ABSTRACT

A self-contained introduction is given to the theory of Gröbner bases which provide algorithmic solutions to many problems in polynomial ideal theory. After explaining the basic theory of Gröbner bases from a term rewriting point of view, we show that abandoning the usual distributive normal form representation of polynomials leads to a considerable simplification of the theory.

1991 Mathematics Subject Classification: 13P10, 68Q40, 68Q42
1987 CR Categories: F.4.1, F.4.2, I.1.2
Key Words and Phrases: Gröbner bases, Buchberger's algorithm, term rewriting systems

[1] Author's address after January 6, 1992: Advanced Research Laboratory, Hitachi Ltd, Hatoyama, Saitama 350-03, Japan; e-mail: ami@harl.hitachi.co.jp.

[2] The work of the second author was performed in partial fulfillment of the requirements for the Master's degree in computer science at the Vrije Universiteit, Amsterdam.

Report CS-R9160
ISSN 0169-118X
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

Introduction

The close relationship between the Knuth-Bendix completion procedure and Buchberger's algorithm for constructing Gröbner bases is well-known. Several people have tried to unify these two procedures. The latest attempt that we are aware of, is the approach of Bündgen [2] who shows that Buchberger's algorithm can be viewed as an extension of the Knuth-Bendix completion procedure to associative and commutative theories. In this paper we are less ambitious. Our goal is to explain the theory underlying Buchberger's algorithm from a rewriting point of view. Historically this is unwarranted since the development of Buchberger's algorithm precedes the invention of the Knuth-Bendix completion procedure by something like five years, but by using the rewrite machinery we are better able to indicate the similarities and the differences between polynomial completion and Knuth-Bendix completion.

The paper is in principle self-contained; however, by its presentation it will be especially suited for term rewriters having no prior knowledge of polynomial completion. In this paper we do not discuss applications of Buchberger's algorithm in polynomial ideal theory. An impressive list of such applications can be found in [1].

The paper is organized as follows. In Section 1 we give a short introduction to rewriting and we explain the theory behind the Knuth-Bendix completion procedure. Section 2 contains a description of the basic notions in polynomial ideal theory. Polynomial rewriting is introduced in Section 3. Section 4 is devoted to Buchberger's algorithm for constructing Gröbner bases. The construction of irreducible Gröbner bases is described in Section 5. In Section 6 we give an account of the two critical pair criteria. We do not claim originality of the material presented in Sections 1–6. Most of the results in Sections 2–6 are due to Buchberger. In Section 7 we show that the construction of Gröbner bases can also be based on the abstract approach of Huet to completion modulo some equivalence relation. To the best of our knowledge this observation is new.

1. Preliminaries

This preliminary section consists of two parts. In the first part we present the basic notions of rewriting in a abstract setting. We give an account of multiset orderings and we recall an early result of Dickson which is the key to termination of the polynomial completion procedure. In the second part we introduce term rewriting and we give a short overview of the completion procedure of Knuth and Bendix.

1.1. Abstract Reduction Systems and Orderings

An *abstract reduction system* (ARS for short) is a structure $\mathcal{A} = \langle A, \rightarrow \rangle$ consisting of a set A and a binary relation $\rightarrow$ on A, named *rewrite relation* or *reduction*. We write $a \leftarrow b$ if $b \rightarrow a$. The *transitive-reflexive closure* of $\rightarrow$ is denoted by $\twoheadrightarrow$. So $a \twoheadrightarrow b$ if there exists a finite, possibly empty, sequence of reduction steps $a = a_1 \rightarrow a_2 \rightarrow \cdots \rightarrow a_n = b$. If $a \twoheadrightarrow b$ then we say that a *reduces* to b and call b a *reduct* of a. We also write $b \twoheadleftarrow a$. The *transitive closure* of $\rightarrow$ is denoted by $\rightarrow^+$. The *symmetric closure* of $\rightarrow$ is denoted by $\leftrightarrow$. So $a \leftrightarrow b$ if $a \rightarrow b$ or $b \rightarrow a$. The *transitive-reflexive-symmetric closure* of $\rightarrow$ is denoted by $\twoheadleftrightarrow$. So $a \twoheadleftrightarrow b$ if there exists is a finite, possibly empty, sequence of steps $a = a_1 \leftrightarrow a_2 \leftrightarrow \cdots \leftrightarrow a_n = b$. This equivalence relation generated by $\rightarrow$ is called *convertibility* or *conversion*. Two elements $a, b \in A$ are *joinable*, denoted by $a \downarrow b$, if there exists a $c \in A$ such that $a \twoheadrightarrow c$ and $b \twoheadrightarrow c$. An element $a \in A$ is a

normal form if there is no $b \in A$ such that $a \to b$. The set of normal forms of $\mathcal{A}$ is denoted by $NF(\mathcal{A})$. We say that *a has a normal form* if there exists a normal form $b \in A$ such that $a \twoheadrightarrow b$.

We now introduce some important properties of ARS's. An ARS $\mathcal{A} = \langle A, \to \rangle$ is *strongly normalizing* if there are no infinite reduction sequences $a_1 \to a_2 \to a_3 \to \cdots$ of elements of A. An ARS $\mathcal{A} = \langle A, \to \rangle$ is *confluent* or has the *Church-Rosser* property if $b \downarrow c$ whenever $a \twoheadrightarrow b$ and $a \twoheadrightarrow c$, for all $a, b, c \in A$. A well-known equivalent formulation of confluence states that conversion coincides with joinability. An ARS $\mathcal{A} = \langle A, \to \rangle$ is *locally confluent* or *weakly Church-Rosser* if $b \downarrow c$ whenever $a \to b$ and $a \to c$, for all $a, b, c \in A$. A *complete* ARS is both confluent and strongly normalizing. Each element is a complete ARS has a unique normal form. The above properties specialize to elements in the obvious way. The following result of Newman [8] forms the theoretical basis for the completion procedure of Knuth and Bendix, to be presented shortly.

NEWMAN'S LEMMA. *Every strongly normalizing and locally confluent ARS is confluent.* □

Newman's Lemma can be viewed as a special case of Lemma 1.1 below. The formulation of that lemma requires the notion of well-founded ordering.

A *partial ordering* is a binary relation $>$ on a set A that is transitive (i.e. if $a > b$ and $b > c$ then $a > c$ for all $a, b, c \in A$) and irreflexive (i.e. $a \not> a$ for all $a \in A$. A partial ordering $>$ on a set A is *total* if for all $a, b \in A$ with $a \neq b$ we either have $a > b$ or $b > a$. We call $>$ *well-founded* if there is no infinite descending sequence $a_1 > a_2 > a_3 > \cdots$ of elements of A. Observe that an ARS $\mathcal{A} = \langle A, \to \rangle$ is strongly normalizing if and only if the transitive closure $\to^+$ of $\to$ is a well-founded ordering on A.

Given an ARS $\mathcal{A} = \langle A, \to \rangle$ and a well-founded ordering $>$ on A, we say that *a is connected to b below c* if there exists a conversion $a = a_1 \leftrightarrow \cdots \leftrightarrow a_n = b$ such that $c > a_i$ for $i = 1, \ldots, n$. This will be denoted as $a \twoheadleftarrow\!\!\twoheadrightarrow_{<c} b$. We call $\mathcal{A}$ *connected* (with respect to $>$) if b and c are connected below a whenever $b \leftarrow a \to c$, for all $a, b, c \in A$. Observe that every connected ARS is strongly normalizing.

LEMMA 1.1 (Winkler and Buchberger [9]). *Every connected ARS is confluent.* □

We will present an elegant proof of this lemma using multiset orderings. A *multiset* over a set A is an unordered collection of elements of A in which elements may occur more than once. If $>$ is a partial ordering on A then we can define a partial ordering $\gg$ on finite multisets over this set A as follows: $M_1 \gg M_2$ if M_2 can be obtained from M_1 by replacing some elements of M_1 (at least one) with a finite number of smaller (with respect to $>$) elements of A. We call $\gg$ the *multiset extension* of $>$.

THEOREM 1.2 (Dershowitz and Manna [3]). *The multiset extension of a well-founded ordering is again a well-founded ordering.* □

PROOF OF LEMMA 1.1. Let $\mathcal{A} = \langle A, \to \rangle$ be a connected ARS with respect to some well-founded ordering $>$ on A. We define an ordering $\ggg$ on conversions in $\mathcal{A}$ as follows:

$$a_1 \leftrightarrow \cdots \leftrightarrow a_n \ggg b_1 \leftrightarrow \cdots \leftrightarrow b_m$$

if $[a_1, \ldots, a_n] \gg [b_1, \ldots, b_m]$. According to Theorem 1.2 $\gg$ is a well-founded ordering on the finite multisets over A. Hence $\ggg$ is a well-founded ordering on conversions in $\mathcal{A}$. We will now show that every conversion $a_1 \leftrightarrow \cdots \leftrightarrow a_n$ that is not a 'valley', i.e. a conversion of the form $a_1 \downarrow a_n$, can be transformed into a conversion between a_1 and a_n which is smaller with respect

to $\ggg$. If $a_1 \leftrightarrow \cdots \leftrightarrow a_n$ is not valley then it contains a 'peak' $a_{i-1} \leftarrow a_i \rightarrow a_{i+1}$. By assumption $a_{i-1} \leftrightarrow\!\!\!\!\!\rightarrow_{<a_i} a_{i+1}$. If we replace the peak $a_{i-1} \leftarrow a_i \rightarrow a_{i+1}$ by the conversion $a_{i-1} \leftrightarrow\!\!\!\!\!\rightarrow_{<a_i} a_{i+1}$, we obtain a conversion between a_1 and a_n which is easily seen to be smaller with respect to $\ggg$. Since $\ggg$ is well-founded, repeating this process eventually results in a valley $a_1 \downarrow a_n$. Hence every pair of convertible elements is joinable. Therefore $\mathcal{A}$ is confluent. □

Next we give an account of Dickson's Lemma (Dickson [4]). This lemma plays a crucial role in the termination proofs of polynomial completion procedures.

DEFINITION 1.3. An infinite sequence $n_1, n_2, n_3, \ldots$ of natural numbers is called *increasing* if $n_i \leqslant n_{i+1}$ for all $i \geqslant 1$.

PROPOSITION 1.4. *Every infinite sequence of natural numbers contains an increasing subsequence.*

PROOF. Let $(n_i)_{i\geqslant 1}$ be an infinite sequence of natural numbers. If some natural number occurs infinitely often in this sequence then we clearly have an increasing subsequence. So suppose that every natural number occurs a finite number of times in the sequence $(n_i)_{i\geqslant 1}$. There are only finitely many numbers in this sequence less than n_1. Hence there exists an index N such that all numbers in the subsequence $(n_i)_{i\geqslant N}$ are greater than or equal to n_1. We now repeat this process with the sequence $(n_i)_{i\geqslant N}$ and eventually we arrive at an increasing subsequence of the original sequence $(n_i)_{i\geqslant 1}$. □

DICKSON'S LEMMA. *If $e_1, e_2, e_3, \ldots$ is an infinite sequence of n-tuples of natural numbers then there exist indices i, j with $i < j$ such that $e_i = (a_1, \ldots, a_n)$, $e_j = (b_1, \ldots, b_n)$ and $a_k \leqslant b_k$ for every $k \in \{1, \ldots, n\}$.*

PROOF. Let us write $(a_1, \ldots, a_n) \lhd (b_1, \ldots, b_n)$ if $a_k \leqslant b_k$ for every $k \in \{1, \ldots, n\}$. By induction on n we will show the existence of an infinite subsequence $e_{i_1} \lhd e_{i_2} \lhd e_{i_3} \lhd \cdots$. The case $n = 1$ has been established in Proposition 1.4. Suppose $e_1, e_2, e_3, \ldots$ is an infinite sequence of $n+1$-tuples. Let $e_i = (a^i_1, \ldots, a^i_{n+1})$ and define $e'_i = (a^i_2, \ldots, a^i_{n+1})$. According to Proposition 1.4 the infinite sequence $(a^i_1)_{i\geqslant 1}$ of first coordinates contains an increasing subsequence $(a^{j_i}_1)_{i\geqslant 1}$. So the sequence $e'_{j_1}, e'_{j_2}, e'_{j_3}, \ldots$ of n-tuples is infinite and hence we obtain an infinite subsequence $e'_{k_1} \lhd e'_{k_2} \lhd e'_{k_3} \lhd \cdots$ from the induction hypothesis. By construction we have also $e_{k_1} \lhd e_{k_2} \lhd e_{k_3} \lhd \cdots$. □

Dickson's Lemma is a special case of Kruskal's Tree Theorem, which forms the theoretical foundation of several well-known methods for proving strong normalization of term rewriting systems.

1.2. Term Rewriting Systems

A *signature* or *alphabet* is a set $\mathcal{F}$ of function symbols. Associated with every function symbol is a natural number denoting its *arity*. Function symbols of arity 0 are called *constants*. The set of *terms* $\mathcal{T}(\mathcal{F}, \mathcal{V})$ built from a signature $\mathcal{F}$ and a countably infinite set of *variables* $\mathcal{V}$ with $\mathcal{F} \cap \mathcal{V} = \varnothing$, is the smallest set containing $\mathcal{V}$ such that $F(t_1, \ldots, t_n) \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ whenever $F \in \mathcal{F}$ has arity n and $t_1, \ldots, t_n \in \mathcal{T}(\mathcal{F}, \mathcal{V})$.

A *term rewriting system* (TRS for short) is a pair $(\mathcal{F}, \mathcal{R})$ consisting of a signature $\mathcal{F}$ and a set $\mathcal{R}$ of *rewrite rules* or *reduction rules*. Every rewrite rule has the form $l \rightarrow r$ with $l, r \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ satisfying the following two constraints:

- the left-hand side l is not a variable,
- the variables which occur in the right-hand side r also occur in l.

In order to define the rewrite relation associated with a given TRS, we first introduce substitutions and contexts.

A *substitution* σ is a mapping from $\mathcal{V}$ to $\mathcal{T}(\mathcal{F},\mathcal{V})$. Substitutions are extended to morphisms from $\mathcal{T}(\mathcal{F},\mathcal{V})$ to $\mathcal{T}(\mathcal{F},\mathcal{V})$, i.e. $\sigma(F(t_1,\ldots,t_n)) = F(\sigma(t_1),\ldots,\sigma(t_n))$ for every n-ary function symbol F and terms $t_1,\ldots,t_n$. We call $\sigma(t)$ an *instance* of t. We write t^σ instead of $\sigma(t)$. An instance of a left-hand side of a rewrite rule is a *redex* (reducible expression).

A *context* $C[\ ]$ is a 'term' which contains precisely one occurrence of a special constant $\Box$. If $C[\ ]$ is a context and t a term then $C[t]$ denotes the result of replacing $\Box$ by t. A term s is a *subterm* of a term t if there exists a context $C[\ ]$ such that $t = C[s]$.

The rewrite rules of a TRS $(\mathcal{F},\mathcal{R})$ define a *rewrite relation* $\rightarrow_\mathcal{R}$ on $\mathcal{T}(\mathcal{F},\mathcal{V})$ as follows: $s \rightarrow_\mathcal{R} t$ if there exists a rewrite rule $l \rightarrow r$ in $\mathcal{R}$, a substitution σ and a context $C[\ ]$ such that $s = C[l^\sigma]$ and $t = C[r^\sigma]$. We say that s rewrites to t by *contracting* redex l^σ. We call $s \rightarrow_\mathcal{R} t$ a *rewrite step* or *reduction step*.

By associating with every TRS $(\mathcal{F},\mathcal{R})$ the ARS $\langle \mathcal{T}(\mathcal{F},\mathcal{V}), \rightarrow_\mathcal{R}\rangle$, all notions defined for ARS's carry over to TRS's. Finite and complete TRS's are of special interest since they have a decidable convertibility relation. The Knuth-Bendix completion procedure attempts to transform a given strongly normalizing TRS into a complete TRS defining the same conversion. We already observed (Newman's Lemma) that it suffices to aim at local confluence.

Let $l_1 \rightarrow r_1$ and $l_2 \rightarrow r_2$ be renamed versions of rewrite rules of a TRS $\mathcal{R}$ such that they have no variables in common. Suppose $l_1 = C[t]$ with $t \notin \mathcal{V}$ such that t and l_2 are unifiable, i.e. $t^\sigma = l_2^\sigma$ for a most general unifier σ. The term $l_1^\sigma = C[l_2]^\sigma$ is subject to the reduction steps $l_1^\sigma \rightarrow r_1^\sigma$ and $l_1^\sigma \rightarrow C[r_2]^\sigma$. The pair of reducts $\langle C[r_2]^\sigma, r_1^\sigma\rangle$ is a *critical pair* of $\mathcal{R}$. If $l_1 \rightarrow r_1$ and $l_2 \rightarrow r_2$ are renamed versions of the same rewrite rule, we do not consider the case $C[\ ] = \Box$. A critical pair $\langle s,t\rangle$ of a TRS $\mathcal{R}$ is *convergent* if $s \downarrow_\mathcal{R} t$. The set of all critical pairs of $\mathcal{R}$ is denoted by $CP(\mathcal{R})$. Furthermore, if $\mathcal{R}_1$ and $\mathcal{R}_2$ are TRS's then $CP(\mathcal{R}_1,\mathcal{R}_2)$ denotes the set of all critical pairs between rules of $\mathcal{R}_1$ and rules of $\mathcal{R}_2$. The following lemma of Huet [5] expresses the significance of critical pairs.

CRITICAL PAIR LEMMA. *A TRS $\mathcal{R}$ is locally confluent if and only if all its critical pairs are convergent.* $\Box$

The basic idea underlying the Knuth-Bendix completion procedure (Knuth and Bendix [6]) is to add a new rewrite rule whenever a non-convergent critical pair is encountered, in order to make it convergent. This has to be repeated until all critical pairs are convergent. In Figure 1 a simple version of the Knuth-Bendix completion procedure is presented. The algorithm presupposes a so-called reduction ordering in order to solve the orientation problem of new rewrite rules in a uniform way.

DEFINITION 1.5.

- A *reduction ordering* $\succ$ is a well-founded ordering on terms which is closed under substitutions and contexts, i.e. if $s \succ t$ then $s^\sigma \succ t^\sigma$ for all substitutions σ and $C[s] \succ C[t]$ for all contexts $C[\ ]$.
- A TRS $\mathcal{R}$ is *compatible* with a reduction ordering $\succ$ if $l \succ r$ for every rewrite rule $l \rightarrow r \in \mathcal{R}$.

It is not difficult to show that a TRS $\mathcal{R}$ is strongly normalizing if and only if there exists a reduction ordering that is compatible with $\mathcal{R}$. The program of Figure 1 has three possibilities:

- it may terminate successfully,
- it may loop infinitely,
- it may fail because a pair of terms cannot be oriented.

Knuth-Bendix completion algorithm: simple version

Input: • a TRS $\mathcal{R}$
• a reduction ordering $\succ$ such that $\mathcal{R}$ is compatible with $\succ$
Output: • a complete TRS $\mathcal{R}'$ with the same conversion as $\mathcal{R}$

$C := CP(\mathcal{R})$;
$\mathcal{R}' := \mathcal{R}$;
while $C \neq \varnothing$ **do**
    choose a pair $\langle s, t\rangle \in C$;
    $C := C - \{\langle s, t\rangle\}$;
    reduce s and t to normal forms s' and t' with respect to $\mathcal{R}'$;
    if $s' \neq t'$ **then**
        if $s' \succ t'$ **then**
            $\alpha := s'$; $\beta := t'$
        else if $t' \succ s'$ **then**
            $\alpha := t'$; $\beta := s'$
        else
            failure
        fi;
        $\mathcal{R}' := \mathcal{R} \cup \{\alpha \to \beta\}$;
        $C := C \cup CP(\mathcal{R}', \{\alpha \to \beta\})$
    fi
od

FIGURE 1.

This is in sharp contrast with polynomial completion procedures which always terminate successfully. In the program of Figure 1 no attempts are made to simplify rewrite rules or to remove redundant rules. Performing such simplifications during the completion process greatly increases efficiency. The completion algorithm of Figure 2 simplifies the rewrite rules as much as possible. Notice that simplifications of left-hand sides and right-hand sides of rewrite rules are treated differently. The algorithm can be made even more efficient by incorporating various *critical pair criteria* which state that certain critical pairs are superfluous. Upon successful termination, the algorithm of Figure 2 delivers a 'fully simplified' TRS.

DEFINITION 1.6. A TRS $\mathcal{R}$ is called *irreducible* or *reduced* if every rewrite rule $l \to r \in \mathcal{R}$ satisfies the following two properties:
(1) l is a normal form with respect to $\mathcal{R} - \{l \to r\}$,
(2) r is a normal form with respect to $\mathcal{R}$.
Observe that a strongly normalizing TRS $\mathcal{R}$ is irreducible if and only if both l and r are normal forms with respect to $\mathcal{R} - \{l \to r\}$, for all rewrite rules $l \to r \in \mathcal{R}$.

According to the following theorem, the result of a successful execution of the simple completion procedure of Figure 1 can always be made irreducible.

THEOREM 1.7 (Métivier [7]). *Every complete TRS can be transformed into an irreducible complete TRS with the same conversion.* □

We conclude this preliminary section with a result that states a kind of unicity for irreducible and complete TRS's.

Knuth-Bendix completion algorithm: efficient version

Input: • a TRS $\mathcal{R}$
• a reduction ordering $\succ$
Output: • a complete irreducible TRS $\mathcal{R}'$ with the same conversion as $\mathcal{R}$

$C := \{\langle l, r\rangle \mid l \to r \in \mathcal{R}\}$;
$\mathcal{R}' := \varnothing$;
while $C \neq \varnothing$ **do**
    choose a pair $\langle s, t\rangle \in C$;
    $C := C - \{\langle s, t\rangle\}$;
    reduce s and t to normal forms s' and t' with respect to $\mathcal{R}'$;
    if $s' \neq t'$ **then**
        if $s' \succ t'$ **then**
            $\alpha := s'$; $\beta := t'$
        else if $t' \succ s'$ **then**
            $\alpha := t'$; $\beta := s'$
        else
            failure
        fi;
        $\mathcal{R}'' := \mathcal{R}' \cup \{\alpha \to \beta\}$;
        for all $l \to r \in \mathcal{R}'$ **do**
            $\mathcal{R}'' := \mathcal{R}'' - \{l \to r\}$;
            reduce l and r to normal forms l' and r' with respect to $\mathcal{R}''$;
            if $l = l'$ **then**
                $\mathcal{R}'' := \mathcal{R}'' \cup \{l' \to r'\}$
            else
                $C := C' \cup \{\langle l', r'\rangle\}$
            fi
        od;
        $\mathcal{R}' := \mathcal{R}''$;
        $C := C \cup CP(\mathcal{R}', \{\alpha \to \beta\})$
    fi
od

FIGURE 2.

THEOREM 1.8 (Métivier [7]). *Let $\mathcal{R}_1$ and $\mathcal{R}_2$ be irreducible complete TRS's with the same conversion. If both TRS's are compatible with a given reduction ordering then they are identical (modulo a renaming of variables in the rewrite rules).* □

2. Polynomial Ideal Theory

In this section we describe the domain in which Buchberger's algorithm operates. In the following we will be working in the ring $K[x_1, \ldots, x_n]$ of *n-variate polynomials* over K. Here K is any *field* and $x_1, \ldots, x_n$ are *indeterminates*. In examples we will use the ring $\mathbb{Q}[x, y, z]$.

DEFINITION 2.1. Let $F \subseteq K[x_1, \ldots, x_n]$ be a finite set of polynomials.
- The *ideal* generated by F is defined as follows:

$$Ideal(F) = \left\{ \sum_{i=1}^{m} h_i f_i \;\middle|\; h_i \in K[x_1, \ldots, x_n] \text{ and } f_i \in F \right\}.$$

- Two polynomials f, g are *congruent modulo* F, notation $f \equiv_F g$, if $f - g \in Ideal(F)$.

In the next two sections we will show that congruence modulo F is decidable, for any finite set F of polynomials.

DEFINITION 2.2. A *power product* is a polynomial of the form $x_1^{i_1} \cdots x_n^{i_n}$. We say that x_j has degree i_j in $x_1^{i_1} \cdots x_n^{i_n}$. The power product $x_1^0 \cdots x_n^0$ is denoted by 1. The set of all power products is denoted by P. A *monomial* is a polynomial of the form $a \cdot p$ with $a \in K$ and $p \in P$. The set of all monomials is denoted by M.

We adopt the usual *distributive normal form* representation of polynomials. This means that every polynomial is a finite sum of monomials whose power products are pairwise distinct. All forthcoming definitions are to be understood with regard to this representation. Only in Section 7 we take a different viewpoint of polynomials.

In the next section we introduce a notion of polynomial reduction. This notion depends on a suitable ordering on power products.

DEFINITION 2.3. An *admissible* ordering $\succ$ is any total ordering on P with the following properties:

- $p \succ 1$ for all $p \in P - \{1\}$,
- if $p_1 \succ p_2$ then $p \cdot p_1 \succ p \cdot p_2$ for all $p, p_1, p_2 \in P$.

Examples of admissible orderings are the *purely lexicographical ordering* and the *total degree ordering*. These are illustrated below.

EXAMPLE 2.4. In the purely lexicographical ordering $\succ_l$ power products are first compared according to the degree of indeterminate x. So $x^2z \succ_l xy^6z^3$. If the degree of x in two power products is the same, then they are compared according to the degree of y. If the degree of y in both power products is also the same, then the power products are ordered according to the degree of z. For example

$$x^3 \succ_l x^2y^2z \succ_l x_2z^2 \succ_l x \succ_l y^3z \succ_l y^2z^2 \succ_l z^5.$$

In the total degree ordering $\succ_t$ power products are ordered according to the sum of the degrees of the indeterminates. If these sums are equal then the purely lexicographical ordering applies. So

$$z^3 \succ_t x^2z \succ_t xy^2 \succ_t xyz \succ_t x^2 \succ_t y^2 \succ_t yz \succ_t x \succ_t 1.$$

DEFINITION 2.5. A power product p_1 is a *divisor* of power product p_2, denoted by $p_1 \lhd p_2$, if there exists a power product p such that $p_1 \cdot p = p_2$.

LEMMA 2.6. *If $p_1, p_2, p_3, \ldots$ is an infinite sequence of power products then there exists indices i, j with $i < j$ such that $p_i \lhd p_j$.*

PROOF. With every power product p_i we associate the n-tuple $e_i = (i_1, \ldots, i_n)$ where i_j is the degree of the indeterminate x_j in p_i. Now we have an infinite sequence $e_1, e_2, e_3, \ldots$ of n-tuples of natural numbers. According to Dickson's Lemma there exists indices i, j with $i < j$ such that $i_k \leqslant j_k$ for $k = 1, \ldots, n$. Hence $p_i \lhd p_j$ since $p_i \cdot p = p_j$ for $p = x_1^{j_1 - i_1} \cdots x_n^{j_n - i_n}$. □

THEOREM 2.7. *Every admissible ordering $\succ$ is well-founded.*

PROOF. Suppose there exists an infinite descending chain $p_1 \succ p_2 \succ p_3 \succ \cdots$ of power products. According to Lemma 2.6 we have $p_i \lhd p_j$ for some $i < j$. Notice that by transitivity $p_i \succ p_j$. We distinguish two cases:

(1) If $p = 1$ then $p_i = p_j$ which contradicts $p_i \succ p_j$.

(2) If $p \neq 1$ then $p \succ 1$ since $\succ$ is admissible. Because an admissible ordering is closed under multiplication, we obtain $p_j = p_i \cdot p \succ p_i \cdot 1 = p_i$ which also contradicts $p_i \succ p_j$.

□

In the remainder of this section we introduce some useful concepts and notations.

DEFINITION 2.8.

- The set of power products occurring in a polynomial t is denoted by $P(t)$ and $M(t)$ denotes the set of monomials occurring in t.
- The *coefficient* of a monomial m is denoted by $\langle m \rangle$ and $\overline{m}$ denotes the remaining power product, so $m = \langle m \rangle \cdot \overline{m}$
- The *least common multiple* of two power products p_1, p_2 is denoted by $lcm(p_1, p_2)$, i.e. $lcm(p_1, p_2)$ is the power product p such that the degree of indeterminate x_i in p equals the maximum of the degrees of x_i in p_1 and p_2. The least common multiple of two monomials is defined as the least common multiple of their power products, i.e. $lcm(m_1, m_2) = lcm(\overline{m_1}, \overline{m_2})$.

DEFINITION 2.9. Let $\succ_P$ be an admissible ordering. The *leading power product* $lp(t)$ of a polynomial $t \neq 0$ is the maximum element in $P(t)$ with respect to $\succ_P$. The *leading monomial* $lm(t)$ of t is the unique monomial in $M(t)$ satisfying $\overline{lm(t)} = lp(t)$. The *leading coefficient* $lc(t)$ of t is the coefficient of $lm(t)$. So $lm(t) = lc(t) \cdot lp(t)$. Finally, $rm(t)$ denotes the polynomial $t - lm(t)$.

EXAMPLE 2.10. Let $t = 3x^2y + 2y^2 - x$. We have $P(t) = \{x^2y, y^2, x\}$ and $M(t) = \{3x^2y, 2y^2, -x\}$. Furthermore, $lp(t) = x^2y$, $lm(t) = 3x^2y$, $lc(t) = 3$ and $rm(t) = 2y^2 - x$, both with respect to the purely lexicographical ordering and the total degree ordering. Let $m_1 = y^2$ and $m_2 = 2x^3y$. We have $\langle m_1 \rangle = 1$, $\overline{m_2} = x^3y$ and $lcm(m_1, m_2) = x^3y^2$.

PROPOSITION 2.11.

(1) $P(s + t) \subseteq P(s) \cup P(t)$.

(2) *If* $P(s) \cap P(t) = \varnothing$ *then* $P(s + t) = P(s) \cup P(t)$.

PROOF.

(1) Trivial.

(2) Since we already know that $P(s + t) \subseteq P(s) \cup P(t)$, it suffices to show that $P(s) \cup P(t) \subseteq P(s + t)$. Let $p \in P(s) \cup P(t)$. From the assumption $P(s) \cap P(t) = \varnothing$ we learn that either $p \in P(s)$ or $p \in P(t)$ and hence $p \in P(s + t)$.

□

3. Polynomial Rewriting

In this section we present a notion of reduction for polynomials and its basic properties. In the next section this polynomial rewrite relation is subjected to a procedure similar to the Knuth-Bendix completion procedure. The ensuing Gröbner bases provide algorithmic solutions to many problems in polynomial ideal theory.

DEFINITION 3.1. A *polynomial rewrite system* (PRS for short) is a pair $(F, \succ_P)$ consisting of a finite set F of polynomials not containing 0 and an admissible ordering $\succ_P$. With every $f \in F$ we associate the *polynomial rewrite rule*

$$f_{\rightarrow}: \quad lm(f) \rightarrow -rm(f).$$

The set of all polynomial rewrite rules associated with F is denoted by $F_{\rightarrow}$. These polynomial rewrite rules induce a rewrite relation $\rightarrow_F$ as follows: $s \rightarrow_F t$ if there exist a monomial $m \in M(s)$, a polynomial rewrite rule $l \rightarrow r \in F_{\rightarrow}$ and a monomial m' such that $m = m'l$ and $t = s - m + m'r$. Occasionally we write $s \rightarrow_F^m t$ to indicate the contracted monomial m. When no confusion can arise we omit the subscript F.

Given the ordering $\succ_P$, F and $F_{\rightarrow}$ can always be constructed from each other. For that reason we will use F and $F_{\rightarrow}$ indifferently. However, in certain cases the use of F is preferred as it leads to more concise formulations. On the other hand, we employ $F_{\rightarrow}$ whenever a concept is introduced that resembles a similar concept in term rewriting. In the following we assume that $\succ_P$ is a fixed admissible ordering and we simply call F a PRS. In examples we will always use the total degree ordering, unless stated otherwise.

By associating with every PRS F the ARS $\langle K[x_1, \ldots, x_n], \rightarrow_F \rangle$, all notions defined in Section 1.1 carry over to polynomial rewriting.

EXAMPLE 3.2. Consider the PRS $F = \{f_1, f_2\}$ with $f_1 = 2x^2y - x^2 - 2$ and $f_2 = 3y^2 - xy + 3x$. The corresponding polynomial rewrite rules are

$$F_{\rightarrow} = \begin{cases} 2x^2y & \rightarrow & x^2 + 2 \\ 3y^2 & \rightarrow & xy - 3x. \end{cases}$$

Consider the polynomial $t = 6x^2y^2 - y^2$. Since $6x^2y^2 = 3y \cdot 2x^2y$, t reduces to $3y \cdot (x^2+2) - y^2 = 3x^2y - y^2 + 6y$ by using the first polynomial rewrite rule. The second rule can be applied in two different ways to t as y^2 divides both x^2y^2 and y^2. Figure 3 shows all possible reduction sequences starting at the polynomial t.

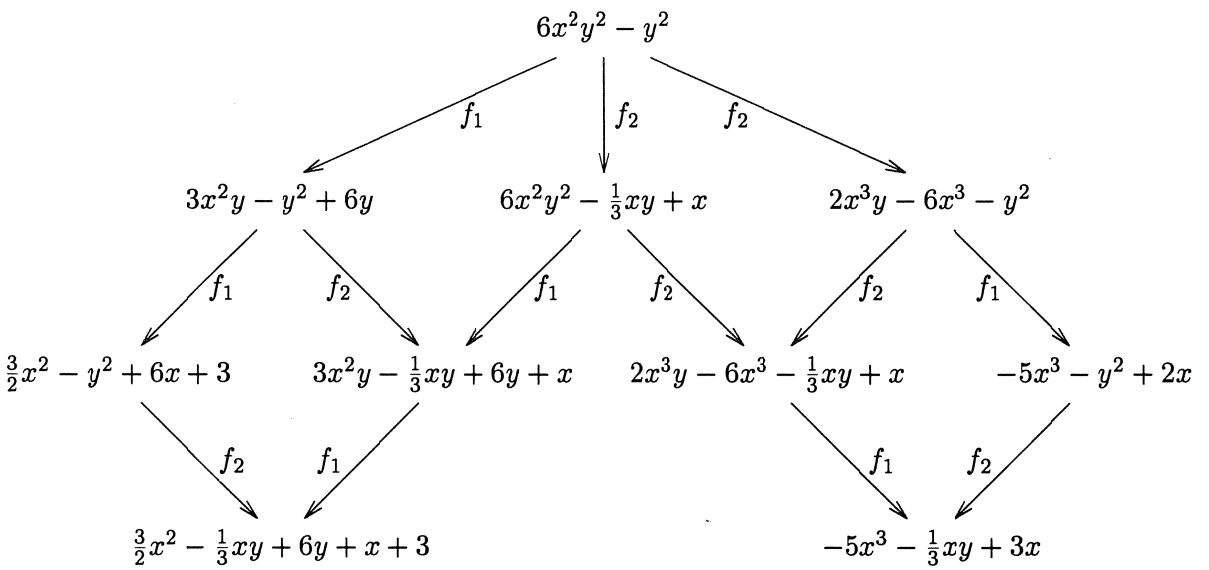

FIGURE 3.

In the remainder of this section we give a few elementary properties of polynomial rewriting. Our first goal is to show that congruence ($\equiv_F$) coincides with conversion ($\leftrightarrow\!\!\!\!\rightarrow_F$). This requires a few preliminary results.

PROPOSITION 3.3. *Let F be a PRS.*

(1) *The relation $\rightarrow_F$ is closed under multiplication by monomials, i.e. if $s \rightarrow_F t$ then $m \cdot s \rightarrow_F m \cdot t$ for all $m \in M$, and hence $\leftrightarrow\!\!\!\!\rightarrow_F$ is closed under multiplication by monomials.*

(2) *If $f \in F$ then $f \rightarrow_F 0$ by application of the polynomial rewrite rule $f_{\rightarrow}$.*

PROOF. Routine. □

The main difference between term rewriting and polynomial rewriting is that the polynomial rewrite relation is not closed under contexts, i.e. the implication $s \rightarrow t \Rightarrow s + u \rightarrow t + u$ does not hold. This considerably complicates the theory of Gröbner bases.

EXAMPLE 3.4. Consider the PRS $F = \{x^2 \rightarrow y\}$ and the polynomials $s = 2x^2 + xy$, $t = xy + 2y$ and $u = x^2 - xy$. We have $s \rightarrow t$, $s + u = 3x^2$ and $t + u = x^2 + 2y$, but $3x^2$ only reduces to $3y$. Actually things are not that bad, since $x^2 + 2y$ also reduces to $3y$.

The next proposition shows that the implication $s \rightarrow t \Rightarrow s + u \downarrow t + u$—which is called *semi-compatibility* in the literature—holds for all polynomials s, t and u (and all PRS's).

PROPOSITION 3.5. *Let F be a PRS.*

(1) *If $s \rightarrow^m t$ and $\overline{m} \notin P(u)$ then $s + u \rightarrow^m t + u$.*

(2) *If $s \rightarrow t$ then $s + u \downarrow t + u$.*

(3) *If $s \leftrightarrow\!\!\!\!\rightarrow t$ then $s + u \leftrightarrow\!\!\!\!\rightarrow t + u$.*

PROOF.

(1) Because $m \in M(s + u)$ this is an immediate consequence of the definition of $\rightarrow$.

(2) By definition there exist a polynomial rewrite rule $l \rightarrow r$ and monomials $m \in M(s)$ and m' such that $m = m'l$ and $t = s - m + m'r$. The case $\overline{m} \notin P(u)$ has been treated in part (1). So assume $\overline{m} \in P(u)$. Let m_1 be the (unique) monomial in u such that $\overline{m_1} = \overline{m}$. We have

$$m_1 = \frac{\langle m_1 \rangle}{\langle m \rangle} m = \frac{\langle m_1 \rangle}{\langle m \rangle} m'l$$

and therefore

$$u \rightarrow u - m_1 + \frac{\langle m_1 \rangle}{\langle m \rangle} m'r.$$

Because $\overline{m_1} \notin P(t)$ we obtain

$$\begin{aligned} t + u &\rightarrow t + u - m_1 + \frac{\langle m_1 \rangle}{\langle m \rangle} m'r \\ &= s - m + m'r + u - m_1 + \frac{\langle m_1 \rangle}{\langle m \rangle} m'r \\ &= s + u - (m + m_1) + \left[1 + \frac{\langle m_1 \rangle}{\langle m \rangle}\right] m'r \end{aligned}$$

from part (1). If $\langle m_1 \rangle = -\langle m \rangle$ then $m + m_1 = 0$ and

$$1 + \frac{\langle m_1 \rangle}{\langle m \rangle} = 0$$

and therefore $t + u \rightarrow s + u$. Otherwise $m + m_1 \in M(s + u)$ and since

$$m + m_1 = \frac{1 + \langle m_1 \rangle}{\langle m \rangle} m'l$$

we obtain

$$s + u \to s + u - (m + m_1) + \left[1 + \frac{\langle m_1 \rangle}{\langle m \rangle}\right] m'r.$$

So in this case $s + u$ and $t + u$ reduce in a single step to a common reduct.

(3) Straightforward consequence of part (2), using induction on the length of $s \leftrightarrow\!\!\!\!\twoheadrightarrow t$.

□

LEMMA 3.6. *The relations $\equiv_F$ and $\leftrightarrow\!\!\!\!\twoheadrightarrow_F$ coincide for every PRS F.*

PROOF.

$\subseteq$ Let $s \equiv_F t$. By definition

$$s - t = \sum_{i=1}^{m} h_i f_i$$

with $f_1, \ldots, f_m \in F$. Without loss of generality we assume that $h_1, \ldots, h_m$ are monomials. We will establish $s \leftrightarrow\!\!\!\!\twoheadrightarrow_F t$ by induction on m. If $m = 0$ then $s = t$. Suppose

$$s - t = \sum_{i=1}^{m+1} h_i f_i$$

or, stated differently,

$$s - (t + h_{m+1} f_{m+1}) = \sum_{i=1}^{m} h_i f_i.$$

The induction hypothesis yields $s \leftrightarrow\!\!\!\!\twoheadrightarrow_F t + h_{m+1} f_{m+1}$. From Proposition 3.3 we obtain $h_{m+1} f_{m+1} \to_F 0$. Proposition 3.5(2) yields $t + h_{m+1} f_{m+1} \downarrow_F t$ and therefore $s \leftrightarrow\!\!\!\!\twoheadrightarrow_F t$.

$\supseteq$ Suppose $s \to_F t$. It is easy to see that $s - t = m \cdot (l - r)$ for some $m \in M$ and polynomial rewrite rule $l \to r \in F_{\to}$. Since $l - r \in F$ we have $m \cdot (l - r) \in \mathit{Ideal}(F)$. Therefore $s \equiv_F t$. The general case follows by a routine induction argument.

□

COROLLARY 3.7. *Let F and G be PRS's. The following statements are equivalent:*

- *F and G define the same ideal,*
- *F and G have the same conversion.*

□

Next we show that polynomial rewriting always terminates. This is a significant difference with term rewriting.

DEFINITION 3.8. The admissible ordering $\succ_P$ on power products is extended to polynomials as follows: $s \succ t$ if $P(s) \succ\!\!\!\succ_P P(t)$ where $\succ\!\!\!\succ_P$ is the multiset extension of the admissible ordering $\succ_P$ on power products. According to Theorem 1.2 $\succ$ is well-founded. Moreover, it is easy to show that $\succ$ is closed under multiplication by monomials.

EXAMPLE 3.9. Consider the reduction step $s = 2x^3 + x^2y - y^2 \to 2x^3 + xy + 3 = t$ in the PRS $\{x^2 \to xy + 3\}$. We have $P(s) = \{x^3, x^2y, y^2\} \succ\!\!\!\succ_P \{x^3, xy, 1\} = P(t)$ since $x^2y \succ_P xy$ and $x^2y \succ_P 1$. Hence $s \succ t$.

PROPOSITION 3.10. *Let F be a PRS. If $s \to_F t$ then $s \succ t$.*

PROOF. By definition there exists a monomial $m \in M(s)$, a polynomial rewrite rule $l \to r \in F_{\to}$ and a monomial m' such that $m = m'l$ and $t = s - m + m'r$. We have $l \succ r$ by definition of polynomial rewrite rule. Therefore $m = m'l \succ m'r$ and thus $P(m) \succcurlyeq_P P(m'r)$. Since $\overline{m} \notin P(s-m)$ we obtain $P(s) = P(s-m) \cup P(m)$ from Proposition 2.11(2). Proposition 2.11(1) yields $P(t) \subseteq P(s-m) \cup P(m'r)$. Combining these statements yields $P(s) \succcurlyeq_P P(t)$. □

THEOREM 3.11. *Every PRS F is strongly normalizing.*

PROOF. Suppose $\to_F$ is not strongly normalizing. According to Proposition 3.10 there exists an infinite descending chain $t_1 \succ t_2 \succ t_3 \succ \cdots$ of polynomials, contradicting the well-foundedness of $\succ$. □

4. Gröbner Bases

Since PRS's are always strongly normalizing, confluence suffices for the decidability of the convertibility relation and hence for the decidability of congruence.

DEFINITION 4.1. A confluent PRS is called a *Gröbner basis.*

In the literature several equivalent formulations of Gröbner bases are reported. Below we list some of them. The easy equivalence proofs are left to the reader.

THEOREM 4.2. *Let F be a PRS. The following statements are equivalent:*

- *F is a Gröbner basis,*
- *every polynomial t has a unique normal form,*
- *every polynomial $t \in Ideal(F)$ reduces to 0,*
- *0 is the only normal form in $Ideal(F)$.*

□

In this section we will show that every PRS can be transformed into a Gröbner basis defining the same conversion, by means of a procedure akin to the Knuth-Bendix completion procedure. Whereas the Knuth-Bendix completion procedure is based on Newman's Lemma, polynomial completion will be based on Lemma 1.1. Before presenting a simple version of the polynomial completion algorithm, we will prove a suitable Critical Pair Lemma for PRS's (Lemma 4.7 below).

DEFINITION 4.3. Let $l_1 \to r_1$ and $l_2 \to r_2$ be different polynomial rewrite rules. Consider the power product $lcm(l_1, l_2)$. Since $lcm(l_1, l_2) = m_1 l_1 = m_2 l_2$ for certain monomials m_1 and m_2, $lcm(l_1, l_2)$ can be reduced both to $m_1 r_1$ and $m_2 r_2$. The pair $\langle m_1 r_1, m_2 r_2 \rangle$ is called a *critical pair.* We call $\langle m_1 r_1, m_2 r_2 \rangle$ *connected* if $m_1 r_1$ and $m_2 r_2$ are connected below $lcm(l_1, l_2)$. In the following we will identify $\langle m_1 r_1, m_2 r_2 \rangle$ and the pair $\langle m_2 r_2, m_1 r_1 \rangle$ originating from the rules $l_2 \to r_2$ and $l_1 \to r_1$. So a PRS with n rules will have $\binom{n}{2}$ critical pairs. The set of all critical pairs of a PRS F is denoted by $CP(F)$ and if F_1 and F_2 are PRS's then $CP(F_1, F_2)$ denotes the set of all critical pairs between rules of $(F_1)_{\to}$ and $(F_2)_{\to}$.

NOTATION. We write $s \succcurlyeq t$ if $s \succ t$ or $P(s) = P(t)$.

PROPOSITION 4.4. *If $s_1 \succ t_1$, $s_2 \succcurlyeq t_2$ and $P(s_1) \cap P(s_2) = \varnothing$ then $s_1 + s_2 \succ t_1 + t_2$.*

PROOF. Straightforward consequence of Proposition 2.11(2). □

The following technical proposition is used in the proof of the Critical Pair Lemma for PRS's, which states that a PRS is a Gröbner basis if and only if all its critical pairs are connected.

PROPOSITION 4.5. *Let F be a PRS.*

(1) *If $s \to^{m_1} t_1$ and $s \to^{m_2} t_2$ with $m_1 \neq m_2$ then t_1 and t_2 can be connected below s.*

(2) *Suppose t_1 and t_2 are connected below s. If $P(s) \cap P(u) = \varnothing$ then $t_1 + u$ and $t_2 + u$ can be connected below $s + u$.*

PROOF.

(1) Let $u = s - m_1 - m_2$. We have $t_1 = n_1 + m_2 + u$ and $t_2 = m_1 + n_2 + u$ for some polynomials n_1, n_2 with $m_1 \to n_1$ and $m_2 \to n_2$. Let $t_3 = n_1 + n_2 + u$. According to Proposition 3.5(2) we have $t_1 \downarrow t_3$ and $t_3 \downarrow t_2$. Proposition 3.10 yields $m_1 \succ n_1$ and $m_2 \succ n_2$. Since $P(m_1)$, $P(m_2)$ and $P(m)$ are pairwise disjoint, two applications of Proposition 4.4 yields $s = m_1 + m_2 + u \succ n_1 + n_2 + u = t_3$. Hence, using Proposition 3.10, all polynomials in the conversion $t_1 \downarrow t_3 \downarrow t_2$ are smaller than s. Therefore $t_1 \leftrightarrow\!\!\!\!\to_{\prec s} t_2$.

(2) By induction on the length of the conversion $t_1 \leftrightarrow\!\!\!\!\to_{\prec s} t_2$ we will show that $t_1 + u \leftrightarrow\!\!\!\!\to_{\prec s+u} t_2 + u$. The case of zero length follows immediately from Proposition 4.4. Suppose $t_1 \to t_1' \leftrightarrow\!\!\!\!\to_{\prec s} t_2$. (The case $t_1 \leftarrow t_1' \leftrightarrow\!\!\!\!\to_{\prec s} t_2$ is similar.) Applying the induction hypothesis to $t_1' \leftrightarrow\!\!\!\!\to_{\prec s} t_2$ yields $t_1' + u \leftrightarrow\!\!\!\!\to_{\prec s+u} t_2 + u$. From Proposition 3.5(2) we obtain $t_1 + u \downarrow t_1' + u$. We already know that $s + u \succ t_1' + u$ and $s + u \succ t_1 + u$ follows from Proposition 4.4. Hence, as a consequence of Proposition 3.10, $t_1 + u \leftrightarrow\!\!\!\!\to_{\prec s+u} t_1' + u$. Combining this conversion with $t_1' + u \leftrightarrow\!\!\!\!\to_{\prec s+u} t_2 + u$ yields the desired result.

□

The next example shows the necessity of the conditions $m_1 \neq m_2$ and $P(s) \cap P(u) = \varnothing$ in Proposition 4.5.

EXAMPLE 4.6.

(1) Consider the PRS

$$F = \left\{ \begin{array}{lcl} xy & \to & x \\ x & \to & 1. \end{array} \right.$$

The monomial xy reduces both to x and y. If x and y are connected below xy then, according to Lemma 4.7 below, F is a Gröbner basis. However, this contradicts the fact that xy has distinct normal forms y and 1.

(2) Consider the PRS $\{xy \to y^2\}$ and the polynomials $s = xy + 1$, $t_1 = y^2$, $t_2 = xy$ and $u = -xy + x$. We have $s \succ t_1$, $s \succ t_2$ and $t_1 \leftarrow t_2$. Thus t_1 and t_2 are connected below s. Notice that $t_1 + u = -xy + x + y^2$ and $t_2 + u = x$ are not connected below $s + u = x + 1$ as $t_1 + u \succ s + u$. And indeed $P(s) \cap P(u) = \{xy\} \neq \varnothing$.

LEMMA 4.7. *A PRS is a Gröbner basis if and only if all its critical pairs are connected.*

PROOF.

$\Rightarrow$ Trivial.

$\Leftarrow$ Consider a PRS F with the property that all its critical pairs are connected. According to Lemma 1.1 it is sufficient to show that F is connected. Let $s \to^{m_1} t_1$ and $s \to^{m_2} t_2$. by application of the polynomial rewrite rules $l_1 \to r_1$ and $l_2 \to r_2$. If $m_1 \neq m_2$ then the result follows from Proposition 4.5(1). Suppose m_1 and m_2 are the same monomial m. If the applied rules are the same then clearly $t_1 = t_2$. Suppose $l_1 \to r_1$ and $l_2 \to r_2$ are different polynomial rewrite rules and let $\langle c_1, c_2 \rangle$ be the corresponding critical pair.

We have $c_1 \leftarrow lcm(l_1, l_2) \rightarrow c_2$ and hence $m'c_1 \leftarrow m \rightarrow m'c_2$ where m' is the monomial such that $m = m' \cdot lcm(l_1, l_2)$. By assumption c_1 and c_2 are connected below $lcm(l_1, l_2)$. Proposition 3.3(1) yields $m'c_1 \leftrightarrow\!\!\!\!\rightarrow_{\prec m} m'c_2$. It is not difficult to show that the premises of Proposition 4.5(2) are fulfilled and hence we obtain $t_1 \leftrightarrow\!\!\!\!\rightarrow_{\prec s} t_2$.

□

Figure 4 shows a simple polynomial completion algorithm. Unlike Knuth-Bendix completion, it always terminates successfully.

Buchberger's algorithm: simple version

Input: • a PRS F
Output: • a Gröbner basis G with the same conversion as F

$C := CP(F)$;
$G := F$;
while $C \neq \varnothing$ **do**
    choose a pair $\langle s, t\rangle \in C$;
    $C := C - \{\langle s, t\rangle\}$;
    reduce s and t to normal forms s' and t' with respect to G;
    if $s' \neq t'$ **then**
        $C := C \cup CP(G, \{s' - t'\})$;
        $G := G \cup \{s' - t'\}$
    fi
od

FIGURE 4.

EXAMPLE 4.8. Consider the PRS

$$F = \begin{cases} x^3y & \rightarrow & x^2 + xy \\ xy^2 & \rightarrow & y^2. \end{cases}$$

Figure 5 gives a graphical representation of the completion algorithm of Figure 4 applied to F. The resulting PRS

$$G = \begin{cases} x^3y & \rightarrow & x^2 + xy \\ xy^2 & \rightarrow & y^2 \\ x^2y & \rightarrow & 0 \\ x^2 & \rightarrow & -xy \\ y^2 & \rightarrow & 0 \end{cases}$$

is a Gröbner basis with the same ideal as F.

The correctness proof of the simple polynomial completion algorithm of Figure 4 employs the following proposition.

PROPOSITION 4.9. *Let F be a PRS and $f \neq 0$ a polynomial.*
(1) *If $f \leftrightarrow\!\!\!\!\rightarrow_F g$ and $g \neq 0$ then $\leftrightarrow\!\!\!\!\rightarrow_{F\cup\{f\}} = \leftrightarrow\!\!\!\!\rightarrow_{F\cup\{g\}}$.*
(2) *If $f \leftrightarrow\!\!\!\!\rightarrow_F 0$ then $\leftrightarrow\!\!\!\!\rightarrow_{F\cup\{f\}} = \leftrightarrow\!\!\!\!\rightarrow_F$.*

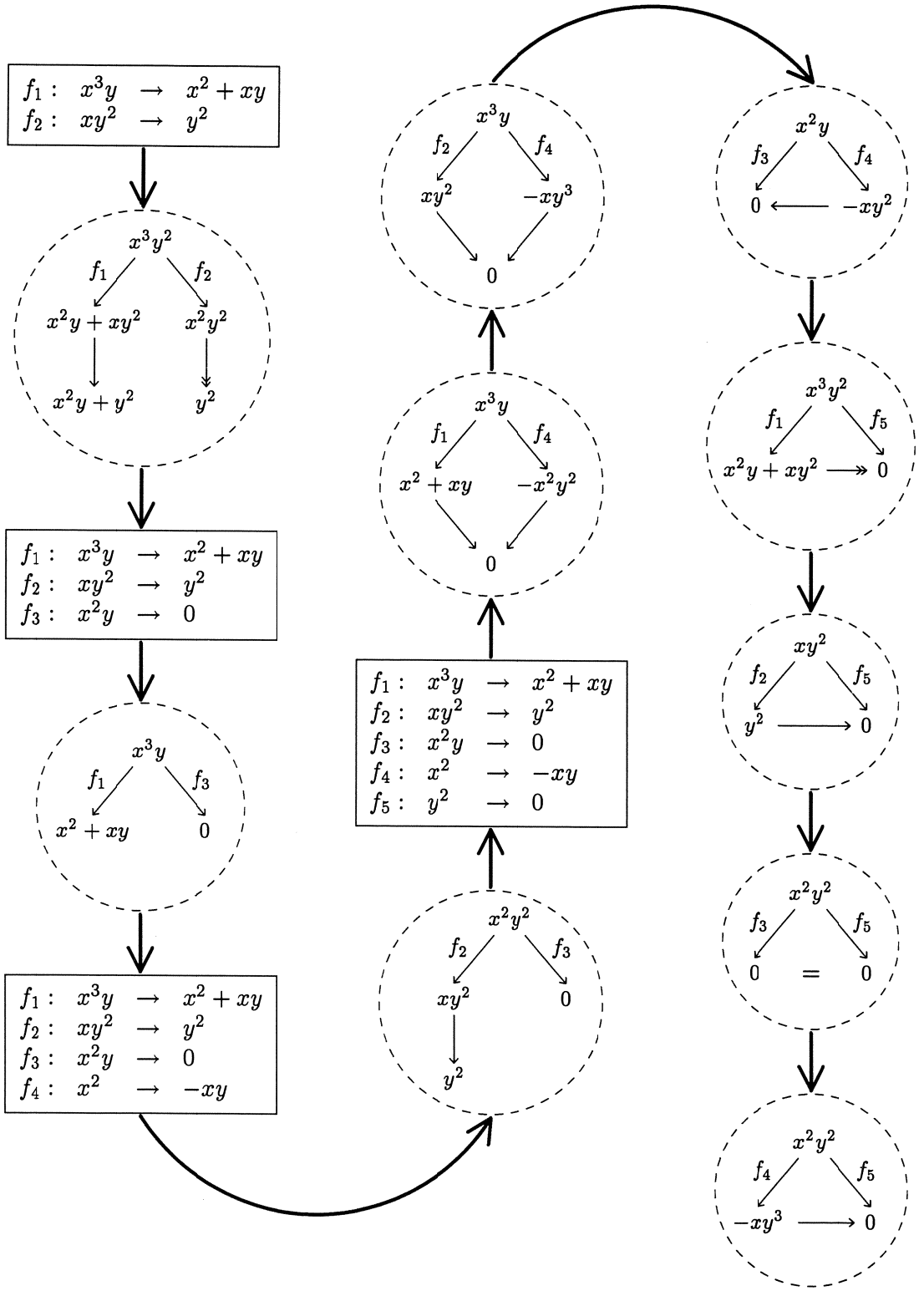
$f_1 : x^3y \to x^2 + xy$
$f_2 : xy^2 \to y^2$
x^3y^2
f_1 f_2
$x^2y + xy^2$ x^2y^2
$x^2y + y^2$ y^2
$f_1 : x^3y \to x^2 + xy$
$f_2 : xy^2 \to y^2$
$f_3 : x^2y \to 0$
x^3y
f_1 f_3
$x^2 + xy$ 0
$f_1 : x^3y \to x^2 + xy$
$f_2 : xy^2 \to y^2$
$f_3 : x^2y \to 0$
$f_4 : x^2 \to -xy$
x^2y^2
f_2 f_3
xy^2 0
y^2
$f_1 : x^3y \to x^2 + xy$
$f_2 : xy^2 \to y^2$
$f_3 : x^2y \to 0$
$f_4 : x^2 \to -xy$
$f_5 : y^2 \to 0$
x^3y
f_1 f_4
$x^2 + xy$ $-x^2y^2$
0
x^3y
f_2 f_4
xy^2 $-xy^3$
0
x^2y
f_3 f_4
0 $-xy^2$
x^3y^2
f_1 f_5
$x^2y + xy^2 \twoheadrightarrow 0$
xy^2
f_2 f_5
$y^2 \to 0$
x^2y^2
f_3 f_5
$0 = 0$
x^2y^2
f_4 f_5
$-xy^3 \to 0$

FIGURE 5.

PROOF.

(1) It suffices to show that $\rightarrow_{\{f\}} \subseteq \leftrightarrow^*_{F\cup\{g\}}$ and $\rightarrow_{\{g\}} \subseteq \leftrightarrow^*_{F\cup\{f\}}$. Suppose $s \rightarrow_{\{f\}} t$. There exists a monomial m such that $t = s - m \cdot f$. Proposition 3.3(1) yields $-m \cdot f \leftrightarrow^*_F -m \cdot g$ and thus $t \leftrightarrow^*_F s - m \cdot g$ by Proposition 3.5(3). Applications of Propositions 3.3(2), 3.3(1) and 3.5(3) yield $s - m \cdot g \leftrightarrow_{\{g\}} s$. Therefore $s \leftrightarrow^*_{F\cup\{g\}} t$. The inclusion $\rightarrow_{\{g\}} \subseteq \leftrightarrow^*_{F\cup\{f\}}$ is almost identical.

(2) Slightly easier than part (1).

□

THEOREM 4.10. *The algorithm of Figure 4 is correct.*

PROOF. We first show the termination of the algorithm. Let f_i be the i-th added polynomial and define $p_i = lp(f_i)$ for $i \geqslant 1$. If the algorithm doesn't terminate then the sequence $p_1, p_2, p_3, \ldots$ must be infinite. Lemma 2.6 yields $p_i \lhd p_j$ for some $i < j$. Since p_i is the *leading* power product of f_i this implies that f_j is reducible by means of the rule $(f_i)_{\rightarrow}$. However, by construction f_j is the difference of two normal forms with respect to all preceding rules. An application of Proposition 2.11(2) shows that f_j is also a normal form with respect to the preceding rules, including $(f_i)_{\rightarrow}$. Thus we obtained a contradiction.

Let C_i and G_i denote the respective values of C and G after the i-th iteration of the while-loop. By induction on i we will show that

(1) $\leftrightarrow^*_{G_i} = \leftrightarrow^*_F$,

(2) if $\langle c_1, c_2\rangle \in CP(G_i) - C_i$ then $c_1 \downarrow_{G_i} c_2$.

The basis of the induction is trivial since $G_0 = F$ and $C_0 = CP(G_0)$. Suppose the statement holds after i iterations of the while-loop and consider iteration $i+1$. We consider two cases:

$s' = t'$ We have $G_{i+1} = G_i$ and $C_{i+1} = C_i - \{\langle s, t\rangle\}$. Property (1) is trivially satisfied. Let $\langle c_1, c_2\rangle \in CP(G_{i+1}) - C_{i+1}$. The case $\langle c_1, c_2\rangle \in CP(G_i) - C_i$ follows from the induction hypothesis. Otherwise, $\langle c_1, c_2\rangle = \langle s, t\rangle$ and $s' = t'$ implies that $s \downarrow_{G_i} t$ and therefore $c_1 \downarrow_{G_{i+1}} c_2$.

$s' \neq t'$ We have $G_{i+1} = G_i \cup \{s'-t'\}$ and $C_{i+1} = C_i - \{\langle s, t\rangle\} \cup CP(G_i, \{s'-t'\})$. By construction, $s' \leftrightarrow^*_{G_i} t'$. Proposition 3.5(3) yields $s'-t' \leftrightarrow^*_{G_i} 0$ and hence, according to Proposition 4.9, the addition of $s'-t'$ doesn't change conversion. It remains to shows that every $\langle c_1, c_2\rangle \in CP(G_{i+1}) - C_{i+1}$ satisfies $c_1 \downarrow_{G_{i+1}} c_2$. The only interesting case is $\langle c_1, c_2\rangle = \langle s, t\rangle$. Since $s' - t' \in G_{i+1}$, we obtain $s' - t' \rightarrow_{G_{i+1}} 0$ from Proposition 3.3. Proposition 3.5(2) yields $s' \downarrow_{G_{i+1}} t'$ and hence also $c_1 \downarrow_{G_{i+1}} c_2$.

Suppose the algorithm terminates after N iterations of the while-loop. Since $C_N = \varnothing$ we obtain the convergence of all critical pairs of G_N from (2). Lemma 4.7 reveals that G_N is a Gröbner basis and property (1) shows that G_N and F have the same conversion. □

5. Irreducible Gröbner Bases

The algorithm of Figure 4 can be optimized by simplifying polynomial rewrite rules during the completion process, similar to the Knuth-Bendix completion procedure. But first we show that these simplifications can also be performed after termination of the simple polynomial completion algorithm.

DEFINITION 5.1. A PRS F is called *irreducible* or *reduced* if every $f \in F$ is a normal form with respect to $F - \{f\}$ and $lc(f) = 1$.

NOTATION. If $f \neq 0$ then $\hat{f}$ denotes the polynomial $f/lc(f)$.

The following result states that an arbitrary PRS can be transformed into an irreducible PRS with the same conversion. So confluence is not essential for this transformation.

THEOREM 5.2. *Every PRS can be transformed into an irreducible PRS with the same conversion.*

PROOF. We will show that repeatedly applying the transition rules of Figure 6 results in an irreducible PRS with the same conversion al as the initial PRS. (The symbol $\uplus$ denotes disjoint union.) We first show that the transition rules preserve convertibility. This is obvious for the

$\frac{F \uplus \{f\}}{F \cup \{g\}}$	if $f \rightarrow_F g$ and $g \neq 0$
$\frac{F \uplus \{f\}}{F}$	if $f \rightarrow_F 0$
$\frac{F \uplus \{f\}}{F \cup \{\hat{f}\}}$	if $lc(f) \neq 1$

FIGURE 6.

third rule. For the first and second rule this is a consequence of Proposition 4.9. Termination of the algorithm is a straightforward consequence of Proposition 3.10. Clearly a PRS is irreducible if and only if none of the transition rules applies. □

Before we can show that the transition rules of Figure 6 preserve confluence, we have to show that the rules do not increase the set of normal forms.

NOTATION. If F is a PRS then $F\nabla$ denotes an irreducible PRS such that $\leftrightarrow\!\!\!\!\rightarrow_{F\nabla} = \leftrightarrow\!\!\!\!\rightarrow_F$ and $NF(F\nabla) \subseteq NF(F)$.

LEMMA 5.3. *Every PRS F can be transformed into a PRS $F\nabla$.*

PROOF. It remains to show that the transition rules of Figure 6 do not increase the set of normal forms. If the third rule is applied then the set of normal forms remains the same. Suppose the first rule is applied. We have to show that $NF(F \cup \{g\}) \subseteq NF(F \cup \{f\})$. Suppose to the contrary that there exists a $t \in NF(F \cup \{g\}) - NF(F \cup \{f\})$. This is only possible if t is reducible by means of $f_{\rightarrow}$. Let $f_{\rightarrow}$ be the rule $l \rightarrow r$. So there exist monomials $m \in M(t)$ and m' such that $m = m'l$. We have $l - r \rightarrow_F g$. From $t \in NF(F)$ we infer that $l \in NF(F)$ and hence $g = l - r'$ for some r' with $r \rightarrow_F r'$. So $g_{\rightarrow}$ is the rule $l \rightarrow r'$. But this is in conflict with the assumption $t \in NF(\{g\})$. We conclude that $NF(F \cup \{g\}) \subseteq NF(F \cup \{f\})$. An application of the second transition rule gives rise to a similar but slightly easier reasoning. □

The next theorem is the analogon of Theorem 1.7 for PRS's. The proof however is much simpler.

THEOREM 5.4. *If F is a Gröbner basis then $F\nabla$ is also a Gröbner basis.*

PROOF. According to Theorem 4.2 it suffices to show that every polynomial has a unique normal form with respect to $F\nabla$. Suppose t has normal forms n_1, n_2 with respect to $F\nabla$. Because F and $F\nabla$ have the same conversion, we obtain $n_1 \leftrightarrow\!\!\!\!\rightarrow_F n_2$ from $n_1 \leftarrow_{F\nabla} t \rightarrow_{F\nabla} n_2$. By definition $NF(F\nabla) \subseteq NF(F)$. Hence $n_1, n_2 \in NF(F)$. Since F is a Gröbner basis we conclude that $n_1 = n_2$. □

EXAMPLE 5.5. Applying the transformation of Theorem 5.2 to the Gröbner basis G of the previous example yields the irreducible Gröbner basis

$$\begin{cases} x^2 & \rightarrow & -xy \\ y^2 & \rightarrow & 0. \end{cases}$$

One might think that every irreducible PRS with the same conversion as a Gröbner basis, is a Gröbner basis. This is not the case. For instance, the irreducible non-confluent PRS

$$\begin{cases} x^2y & \rightarrow & x \\ xy^2 & \rightarrow & x \end{cases}$$

and the reducible Gröbner basis

$$\begin{cases} x^2y & \rightarrow & x \\ xy^2 & \rightarrow & x \\ x^2 & \rightarrow & xy \end{cases}$$

have the same conversion.

In Figure 7 a more efficient polynomial completion algorithm is presented. The simplification of the polynomial rewrite rules takes place in the statements "$G := F\nabla$" and "$G := (G \cup \{n\})\nabla$". The assignment "$C := CP(G)$" in the if-statement is a bit unrealistic since many critical pairs

Buchberger's algorithm: efficient version

Input: • a PRS F
Output: • an irreducible Gröbner basis G with the same conversion as F

$G := F\nabla$;
$C := CP(G)$;
while $C \neq \varnothing$ **do**
    choose a pair $\langle s, t\rangle \in C$;
    reduce $s - t$ to normal form n with respect to G;
    if $n \neq 0$ **then**
        $G := (G \cup \{n\})\nabla$;
        $C := CP(G)$
    else
        $C := C - \{\langle s, t\rangle\}$
    fi
od

FIGURE 7.

of G are already known to be connected. The reason for adopting this simple-minded approach is that the precise bookkeeping of the critical pairs that can be relegated to oblivion and those that have to be computed is rather tedious, see Buchberger [1] (Algorithm 6.3) for details.

EXAMPLE 5.6. Consider again the PRS F of Example 4.8. The application of the completion algorithm of Figure 7 to F is illustrated in Figure 9. In this figure a dotted arrow denotes an application of one of the transition rules of Figure 8. These rules clearly implement the

$\dfrac{F \uplus \{f\}}{F \cup \{g\}}$	if $f \to_F^+ g$, $g \neq 0$ and $g \in NF(F)$
$\dfrac{F \uplus \{f\}}{F}$	if $f \to_F^+ 0$
$\dfrac{F \uplus \{f\}}{F \cup \{\hat{f}\}}$	if $lc(f) \neq 1$

FIGURE 8.

∇-operation. Observe that only two critical pairs are computed compared to the ten pairs in Example 4.8. In the next section we will see that the second pair actually is superfluous.

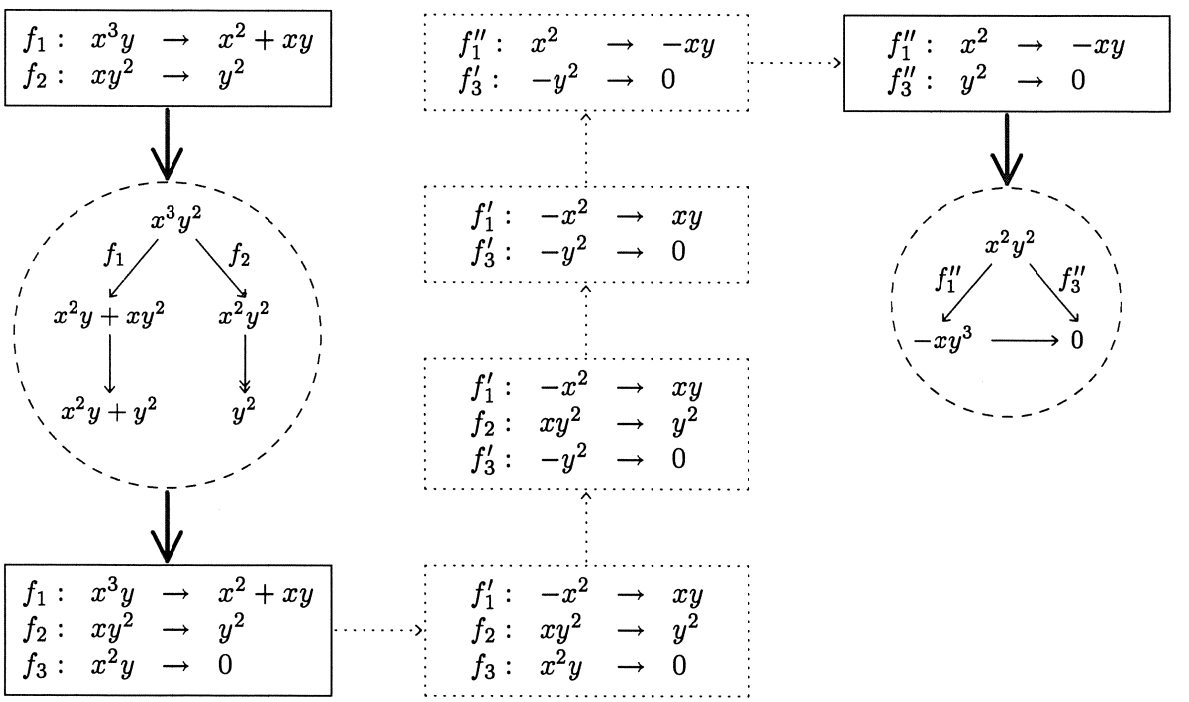

FIGURE 9.

PROPOSITION 5.7. *Let F be a PRS. Suppose t_1 and t_2 are connected below a monomial s. If $s \succ u$ then $t_1 + u$ and $t_2 + u$ are also connected below s.*

PROOF. The proof is similar to that of Proposition 4.5(2). The restriction to monomials s ensures that $s \succ t$ and $s \succ u$ imply $s \succ t + u$, for all polynomials t. □

THEOREM 5.8. *The algorithm of Figure 7 is correct.*

PROOF. We first show the termination of the algorithm. Let f_i be the i-th added polynomial and define $p_i = lp(f_i)$ for $i \geqslant 1$. Suppose the algorithm doesn't terminate. It is easy to see that the sequence $f_1, f_2, \ldots$ is infinite. Define PRS's G^i for $i \geqslant 0$ as follows: $G^0 = F\nabla$ and $G^{i+1} = (G^i \cup \{f^{i+1}\})\nabla$. Using the definition of ∇, a straightforward induction argument shows that

$$NF(G^i) \subseteq NF(\{f_1, \ldots, f_i\})$$

for all $i \geqslant 0$. Lemma 2.6 yields $p_i \lhd p_j$ for some $i < j$. Since p_i is the *leading* power product of f_i this implies that f_j is reducible by means of the rule $(f_i)_{\rightarrow}$. However, by construction $f_j \in NF(G^{j-1})$. The above inclusion shows that $f_j \in NF(\{f_1, \ldots, f_{j-1}\})$. In particular f_j is a normal form with respect to $(f_i)_{\rightarrow}$. Thus we obtained a contradiction.

Let C_i and G_i denote the respective values of C and G after the i-th iteration of the while-loop. By induction on i we will show that

(1) $\leftrightarrow\!\!\!\!\rightarrow_{G_i} = \leftrightarrow\!\!\!\!\rightarrow_F$,

(2) every critical pair in $CP(G_i) - C_i$ is connected (with respect to G_i).

The basis of the induction is trivial since $G_0 = F$ and $C_0 = CP(G_0)$. Suppose the statement holds after i iterations of the while-loop and consider iteration $i + 1$. We consider two cases:

$n = 0$ We have $G_{i+1} = G_i$ and $C_{i+1} = C_i - \{\langle s, t\rangle\}$. Property (1) is trivially satisfied. Let $\langle c_1, c_2\rangle \in CP(G_{i+1}) - C_{i+1}$. The case $\langle c_1, c_2\rangle \in CP(G_i) - C_i$ follows from the induction hypothesis. Otherwise, $\langle c_1, c_2\rangle = \langle s, t\rangle$. Let $l_1 \rightarrow r_1$ and $l_2 \rightarrow r_2$ be the polynomial rewrite rules that gave rise to the pair $\langle c_1, c_2\rangle$. We have $c_1 - c_2 \twoheadrightarrow_{G_i} 0$ and $lcm(l_1, l_2) \succ c_1, c_2$. Since $lcm(l_1, l_2) \succ c_1 - c_2$, $c_1 - c_2$ and 0 are connected below $lcm(l_1, l_2)$ in G_i. An application of Proposition 5.7 shows that c_1 and c_2 can be connected below $lcm(l_1, l_2)$ in $G_i = G_{i+1}$.

$n \neq 0$ We have $G_{i+1} = (G_i \cup \{n\})\nabla$ and $C_{i+1} = CP(G_{i+1})$. By construction, $s \leftrightarrow\!\!\!\!\rightarrow_{G_i} t$ and $s - t \twoheadrightarrow_{G_i} n$. Proposition 3.5(3) yields $n \leftrightarrow\!\!\!\!\rightarrow_{G_i} 0$ and hence, as a consequence of Proposition 4.9, the addition of n doesn't change conversion. Since $CP(G_{i+1}) - C_{i+1} = \varnothing$, property (2) is trivially satisfied.

Suppose the algorithm terminates after N iterations of the while-loop. Since $C_N = \varnothing$ we infer from property (2) that all critical pairs of G_N are connected. Lemma 4.7 reveals that G_N is a Gröbner basis and property (1) shows that G_N and F have the same conversion. □

This section is concluded with a uniqueness result for irreducible Gröbner bases. The proof is similar to that of the analogous result for TRS's (Theorem 1.8).

THEOREM 5.9. *Irreducible Gröbner bases with respect to the same underlying ordering and with the same conversion are identical.*

PROOF. Let F and G be irreducible Gröbner bases such that $\leftrightarrow\!\!\!\!\rightarrow_F = \leftrightarrow\!\!\!\!\rightarrow_G$. Before proving $F = G$, we show that F and G define the same normal forms. Suppose to the contrary that there exists a $t \in NF(F) - NF(G)$. So $t \rightarrow_G t'$ for some polynomial t'. Since $\leftrightarrow\!\!\!\!\rightarrow_F = \leftrightarrow\!\!\!\!\rightarrow_G$ we have $t \leftrightarrow\!\!\!\!\rightarrow_F t'$. Because F is a Gröbner basis and $t \in NF(F)$, this implies $t' \rightarrow_F^+ t$. But now we both have $t \succ t'$ (since $t \rightarrow_G t'$) and $t' \succ t$ (since $t' \rightarrow_F^+ t$), which is impossible. Therefore $NF(F) \subseteq NF(G)$ and, by symmetry, $NF(F) = NF(G)$. In order to show that $F = G$ it suffices to show that $F_{\rightarrow} \subseteq G_{\rightarrow}$. Let $l \rightarrow r \in F_{\rightarrow}$. Since F is irreducible we know that r is a normal form. Using $\leftrightarrow\!\!\!\!\rightarrow_F = \leftrightarrow\!\!\!\!\rightarrow_G$ and the fact that G is a Gröbner basis, we infer $l \rightarrow_G^+ r$. Let $l' \rightarrow r' \in G_{\rightarrow}$ be the rule that is applied first in this sequence. Because $\langle l\rangle = \langle l'\rangle = 1$ we obtain $l = l'$ from the irreducibility of F. Notice that r' is a normal form. Confluence of G yields $r = r'$. Therefore $l \rightarrow r \in G_{\rightarrow}$. □

6. Critical Pair Criteria

In this section we present two criteria which show that certain critical pairs do not need to be considered during the completion process. The first one states that the critical pair between rules whose left-hand sides have no common indeterminates is superfluous. The correctness of this criterion relies on the following proposition.

PROPOSITION 6.1. *Let s_1 be a monomial and s_2, t polynomials. If $s_1 \to s_2$ then $s_1 t \twoheadrightarrow s_2 t$.*

PROOF. If $t = 0$ then we have nothing to prove. If $t \neq 0$ then there exist monomials $m_1 \succ m_2 \succ \cdots \succ m_n$ $(n \geqslant 1)$ such that

$$t = \sum_{i=1}^{n} m_i.$$

Define

$$u_j = \sum_{i=1}^{j-1} s_2\, m_i + \sum_{i=j}^{n} s_1 m_i$$

for $j = 1, \ldots, n+1$. Notice that $u_1 = s_1 t$ and $u_{n+1} = s_2 t$. By induction on j we will show that $s_1 t \twoheadrightarrow u_j$. The case $j = 1$ is trivial. Suppose $s_1 t \twoheadrightarrow u_j$ for some $j \in \{1, \ldots, n\}$. We will prove that $u_j \to u_{j+1}$. The important observation is that $s_1 m_j \in M(u_j)$. This is an immediate consequence of the following two facts, which can be easily proved:

$$\text{if } m \in M(\sum_{i=1}^{j-1} s_2\, m_i) \text{ then } s_1 m_j \succ m,$$

$$\text{if } m \in M(\sum_{i=j}^{n} s_1 m_i) \text{ then } m \succ s_1 m_j.$$

Proposition 3.3(1) yields $s_1 m_j \to s_2\, m_j$ and from Proposition 3.5(1) we infer that

$$u_j = \sum_{i=1}^{j-1} s_2\, m_i + s_1 m_j + \sum_{i=j+1}^{n} s_1 m_i \to \sum_{i=1}^{j-1} s_2\, m_i + s_2\, m_j + \sum_{i=j+1}^{n} s_1 m_i = u_{j+1}.$$

□

EXAMPLE 6.2. Consider the polynomial rewrite rule $s_1 = x \to 1 = s_2$ and the polynomial $t = x + 1$. We have $s_1 t = x^2 + x \to x^2 + 1 \to x + 1 = s_2\, t$. It is essential that we start reducing the least monomial in $s_1 t$, otherwise we do not reach $s_2\, t$: $x^2 + x \to x + x = 2x \to 2 \neq x + 1$.

LEMMA 6.3. *If $l_1 \to r_1$ and $l_2 \to r_2$ are polynomial rewrite rules such that $lcm(l_1, l_2) = \overline{l_1 l_2}$ then their critical pair is convergent.*

PROOF. The critical pair involved is

$$\left\langle \frac{r_1}{\langle l_1 \rangle} \overline{l_2},\ \overline{l_1} \frac{r_2}{\langle l_2 \rangle} \right\rangle.$$

Proposition 6.1 shows that $\dfrac{r_1}{\langle l_1 \rangle} \overline{l_2} \twoheadrightarrow \dfrac{r_1}{\langle l_1 \rangle} \dfrac{r_2}{\langle l_2 \rangle} \twoheadleftarrow \overline{l_1} \dfrac{r_2}{\langle l_2 \rangle}$. □

EXAMPLE 6.4. The criterion of Lemma 6.3 reveals that the critical pair between f_1'': $x^2 \to -xy$ and f_3'': $y^2 \to 0$ in the completion process of Figure 9 is superfluous.

The second critical pair criterion requires a more complicated formulation. Moreover, it is difficult to implement. The optimized version of the completion algorithm in Buchberger [1] (Algorithm 6.3) only detects the case $n = 3$. Fortunately, the correctness proof is straightforward.

LEMMA 6.5. *Suppose we have polynomial rewrite rules $l_i \to r_i$ for $i = 1, \ldots, n$ such that the critical pair between $l_i \to r_i$ and $l_{i+1} \to r_{i+1}$ is connected for $i = 1, \ldots, n-1$. If $l_i \lhd lcm(l_1, l_n)$ for all $i \in \{2, \ldots, n-1\}$ then the critical pair between $l_1 \to r_1$ and $l_n \to r_n$ is connected.*

PROOF. Clearly $l_1 \lhd lcm(l_1, l_n)$ and $l_n \lhd lcm(l_1, l_n)$. Hence for every $i \in \{1, \ldots, n\}$ there exists a monomial m_i such that $m_i l_i = lcm(l_1, l_n)$. Let $i \in \{1, \ldots, n-1\}$. We will show that $m_i r_i$ and $m_{i+1} r_{i+1}$ are connected below $lcm(l_1, l_n)$. By assumption $m' r_i$ and $m'' r_{i+1}$ are connected below $lcm(l_i, l_{i+1})$. Here m' and m'' are the monomials that satisfy $m' l_i = m'' l_{i+1} = lcm(l_i, l_{i+1})$. From $l_i \lhd lcm(l_1, l_n)$ and $l_{i+1} \lhd lcm(l_1, l_n)$ we infer that $lcm(l_i, l_{i+1}) \lhd lcm(l_1, l_n)$. Hence there exists a power product p such that $p \cdot lcm(l_i, l_{i+1}) = lcm(l_1, l_n)$. We have $m'p = m_i$ and $m''p = m_{i+1}$. Repeated application of Proposition 3.3(1) shows that $m_i r_i$ and $m_{i+1} r_{i+1}$ are connected below $lcm(l_1, l_n)$. Therefore $m_1 r_1$ and $m_n r_n$ are connected (via $m_2 r_2, \ldots, m_{n-1} r_{n-1}$) below $lcm(l_1, l_n)$. Since $\langle m_1 r_1, m_n r_n \rangle$ is the critical pair between $l_1 \to r_1$ and $l_n \to r_n$, we are done. □

EXAMPLE 6.6. Consider the PRS

$$F = \left\{ \begin{array}{lcl} x^2 & \to & xy \\ y^2 & \to & y \\ xy^2 & \to & x^2. \end{array} \right.$$

The critical pair between the first two rules is convergent according to Lemma 6.3. The critical pair between the last two rules is also convergent: $xy \leftarrow x^2$. Since $y^2 \lhd x^2y^2 = lcm(x^2, xy^2)$, Lemma 6.5 shows that the critical pair between the first and the last rule is connected. Hence F is a Gröbner basis.

The next example shows why we required $l_i \lhd lcm(l_1, l_n)$ in Lemma 6.5.

EXAMPLE 6.7. Consider the PRS

$$F = \left\{ \begin{array}{lcl} x & \to & 1 \\ y^2 & \to & y \\ xy & \to & x. \end{array} \right.$$

The critical pair between $x \to 1$ and $y^2 \to y$ is convergent according to Lemma 6.3. The critical pair between $y^2 \to y$ and $xy \to x$ is trivially convergent: $xy = xy$. We do not have $y^2 \lhd xy = lcm(x, xy)$ and indeed the critical pair $\langle y, x \rangle$ between $x \to 1$ and $xy \to x$ is not connected because otherwise F would be a Gröbner basis by Lemma 4.7, which is not the case as y and 1 are different normal forms of xy.

7. Efficiency versus Simplicity

In the preceding sections we have seen that the lack of "closure under contexts" of the polynomial rewrite relation is quite an annoyance in the development of the theory of Gröbner bases. In this

section we propose a possible remedy. The basic idea is very simple. We abandon the distributive normal form representation and we 'define' a polynomial as a finite sum of monomials instead. So $x - x + y$, $2y - y$ and y are viewed as different polynomials. Every polynomial t corresponds to a unique polynomial in distributive normal form, which we will denote by $\tilde{t}$.

DEFINITION 7.1. Two polynomials t_1, t_2 are *equivalent*, denoted by $t_1 \sim t_2$, if $\tilde{t}_1 = \tilde{t}_2$.

It is easy to see that $\sim$ is an equivalence relation. Moreover, every equivalence class of polynomials contains precisely one distributive normal form.

The concept of PRS remains unchanged: we allow only polynomial rewrite rules stemming from distributive normal forms. However, the induced rewrite relation does change since the resulting polynomials are no longer put into distributive normal form.

EXAMPLE 7.2. Consider the Gröbner basis

$$F = \begin{cases} x^2 & \to & 0 \\ x & \to & y - 1 \\ y & \to & 1. \end{cases}$$

The following normalizing reduction sequences show that the new polynomial rewrite relation is not confluent:

$$\begin{aligned} x^2 - xy &\to xy - x - xy \\ &\to y^2 - y - x - xy \\ &\to y^2 - y - y + 1 - xy \\ &\to y^2 - y - y + 1 - y^2 + y \\ &\to y - y - y + 1 - y^2 + y \\ &\to y - y - y + 1 - y + y \\ &\twoheadrightarrow 1 - 1 - 1 + 1 - 1 + 1, \end{aligned}$$

$$\begin{aligned} x^2 - xy &\to -xy \\ &\to -x \\ &\to -y + 1 \\ &\to -1 + 1. \end{aligned}$$

Notice that $1 - 1 - 1 + 1 - 1 + 1$ and $-1 + 1$ are equivalent normal forms.

The above example shows that the new polynomial rewrite relation is very inefficient. The example makes also clear that we have to redefine the concept of Gröbner basis since we can only hope for confluence *modulo* $\sim$. As far as theory is concerned, these drawbacks do not counterbalance the profit we obtain from the property of "closure under contexts", which holds by definition so to say. The new completion process will be based on the abstract approach to completion modulo some equivalence as developed in Huet [5].

DEFINITION 7.3. Let $\mathcal{A} = \langle A, \to \rangle$ be an ARS and $\sim$ an equivalence relation on A.

- Two elements $a, b \in A$ are *joinable modulo* $\sim$, denoted by $a \downarrow_\sim b$, if there exist $c, d \in A$ such that $a \twoheadrightarrow c \sim d \twoheadleftarrow d$.

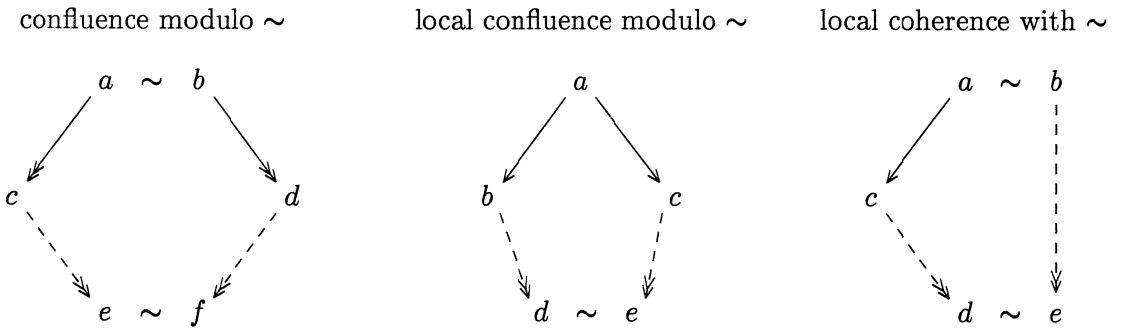

FIGURE 10.

- We say that $\mathcal{A}$ (or $\rightarrow$) is *confluent modulo* $\sim$ if $c \downarrow_\sim d$ whenever $c \twoheadleftarrow a \sim b \twoheadrightarrow d$, for all $a, b, c, d \in A$.
- We call $\mathcal{A}$ is *locally confluent modulo* $\sim$ if $b \downarrow_\sim c$ whenever $b \leftarrow a \rightarrow c$, for all $a, b, c \in A$.
- We call $\mathcal{A}$ is *locally coherent with* $\sim$ if $b \downarrow_\sim c$ whenever $b \leftarrow a \sim c$, for all $a, b, c \in A$.

LEMMA 7.4 (Huet [5]). *Let $\mathcal{A} = \langle A, \rightarrow \rangle$ be an ARS and $\sim$ an equivalence relation on A. If*
- *$\mathcal{A}$ is strongly normalizing,*
- *$\mathcal{A}$ is locally confluent modulo $\sim$, and*
- *$\mathcal{A}$ is locally coherent with $\sim$*

then $\mathcal{A}$ is confluent modulo $\sim$. □

The proof given in [5] is highly combinatorial. We will give a multiset argument. The idea is to associate with every conversion (consisting of $\rightarrow$, $\leftarrow$ and $\sim$-steps) a multiset such that an application of one of the local properties results in a conversion whose associated multiset is smaller with respect to the multiset extension of $\rightarrow^+$. We have to be careful though, since the mapping from conversions to multisets used in Lemma 1.1 doesn't work here. Consider for example the ARS $\{a \rightarrow b,\ a \rightarrow c,\ d \rightarrow e\}$ with equivalences $a \sim d$ and $c \sim e$. Local coherence transforms the conversion $b \leftarrow a \sim d$ into $b \rightarrow c \sim e \leftarrow d$. But the multiset $[b, a, d]$ is not greater than $[b, c, e, d]$ as $a \not\rightarrow^+ e$. The trick is to use a slightly different translation from conversions to multisets:

step in conversion	contribution to multiset
$a \rightarrow b$	a
$a \leftarrow b$	b
$a \sim b$	a, a, b, b

In our example this gives rise to the multisets $[a, a, a, d, d]$ and $[b, c, c, e, e, d]$. Now the former is greater than the latter since $a \rightarrow^+ b$, $a \rightarrow^+ c$ and $d \rightarrow^+ e$.

PROOF OF LEMMA 7.4. The proof has the same structure as that of Lemma 1.1. Conversions consisting of $\rightarrow$, $\leftarrow$ and $\sim$-steps are mapped to multisets according to the above translation scheme. If a conversion is not a 'valley modulo $\sim$', i.e. of the form $a \downarrow_\sim b$, then one of the following situations must occur.
- $a \leftarrow b \rightarrow c$

  An application of local confluence modulo $\sim$ results in a smaller conversion since all new elements are reducts of b.

- $a \leftarrow b \sim c$ or $a \sim b \rightarrow c$
  Local coherence of $\rightarrow$ with $\sim$ results in a multiset which is easily shown to be smaller in the multiset extension of $\rightarrow^+$.
- $a \sim b \sim c$
  Since $\sim$ is transitive, we may replace this situation by $a \sim c$.

Repeating these transformations eventually results in a conversion of the form $a \downarrow_\sim b$. Hence $\mathcal{A}$ is confluent modulo $\sim$. □

Now we redefine the concept of Gröbner basis. A PRS F is called a Gröbner basis if $\rightarrow_F$ is confluent modulo $\sim$. Whereas the previous completion procedures were based on Lemma 1.1, the new representation calls for Lemma 7.4. Two of the three conditions in Lemma 7.4 are always satisfied. Strong normalization remains valid in the new setting. The proof (of Proposition 3.10 that is) even becomes simpler because Proposition 2.11 is no longer needed. The following result shows that local coherence is never a problem.

LEMMA 7.5. *Every PRS F is locally coherent with $\sim$.*

PROOF. Let $s \rightarrow^m t$ and $s \sim u$. Decompose s into $s_1 + s_2$ such that the following equivalence holds for all $m' \in M(s)$: $\overline{m'} = \overline{m}$ if and only if $m' \in M(s_1)$. Likewise we decompose u into $u_1 + u_2$. It is not difficult to see that $s_1 \sim u_1$ and $s_2 \sim u_2$. From $s_1 \sim u_1$ we infer that

$$\sum_{m' \in M(s_1)} \langle m' \rangle \;=\; \sum_{m' \in M(u_1)} \langle m' \rangle .$$

Let us denote this coefficient by c. We may write $s_1 = m + s_3$. By definition $m = m_1 l$ for some monomial m_1 and polynomial rewrite rule $l \rightarrow r$. Therefore

$$t \;=\; \sum_{m_2 \in M(r)} m_1 m_2 + s_3 + s_2 .$$

If $m' \in M(s_3)$ then $m' = \frac{\langle m' \rangle}{\langle m \rangle} m_1 l$. So all monomials in s_3 can be reduced by means of the rule $l \rightarrow r$. Hence

$$\begin{aligned} t \;&\twoheadrightarrow\; \sum_{m_2 \in M(r)} m_1 m_2 + \sum_{m' \in M(s_3)} \sum_{m_2 \in M(r)} \frac{\langle m' \rangle}{\langle m \rangle} m_1 m_2 + s_2 \\ &=\; \underbrace{\sum_{m' \in M(s_1)} \sum_{m_2 \in M(r)} \frac{\langle m' \rangle}{\langle m \rangle} m_1 m_2}_{t'} + s_2 . \end{aligned}$$

Similarly, the reduction of all monomials in u_1 gives rise to the sequence

$$u \;\twoheadrightarrow\; \underbrace{\sum_{m' \in M(u_1)} \sum_{m_2 \in M(r)} \frac{\langle m' \rangle}{\langle m \rangle} m_1 m_2}_{u'} + u_2 .$$

We have $t' \sim u'$ since

$$\hat{t'} \;=\; \sum_{m_2 \in M(r)} \frac{c}{\langle m \rangle} m_1 m_2 \;=\; \hat{u'}$$

Combining this with $s_2 \sim u_2$ yields the desired $t' + s_2 \sim u' + u_2$. □

EXAMPLE 7.6. Let $s = x^2y+2x+2x^2y$, $t = 3x^2-2x+2x+2x^2y$ and $u = 9x^2+x+3x^2y-9x^2+x$. We have $s \sim u$ and $s \to t$ by an application of the polynomial rewrite rule $xy \to 3x - 2$. Reduction of the monomials $2x^2y$ in t and $3x^2y$ in u yields $t \to 3x^2 - 2x + 2x + 6x^2 - 4x$ and $u \to 9x^2 + x + 9x^2 - 6x - 9x^2 + x$. The resulting polynomials are clearly equivalent.

So in order to complete a given PRS into a Gröbner basis, it suffices to fulfill the second requirement of Lemma 7.4: local confluence modulo $\sim$. This property is enforced by the joinability modulo $\sim$ of critical pairs.

DEFINITION 7.7. A pair $\langle s, t\rangle$ is *convergent modulo* $\sim$ if s and t are joinable modulo $\sim$.

LEMMA 7.8. *A PRS is locally confluent modulo* $\sim$ *if and only if all its critical pairs are convergent modulo* $\sim$.

PROOF.

$\Rightarrow$ Trivial.

$\Leftarrow$ Due to "closure under contexts" of the polynomial rewrite relation, the proof is much simpler than that of Lemma 4.7. Consider a PRS with the property that all its critical pairs are convergent modulo $\sim$. Let $s \to^{m_1} t_1$ and $s \to^{m_2} t_2$. by application of the polynomial rewrite rules $l_1 \to r_1$ and $l_2 \to r_2$. We distinguish two cases.

(1) If m_1 and m_2 are different monomial occurrences in s then we clearly have $t_1 \to^{m_2} t_3 \leftarrow^{m_1} t_2$ for some polynomial t_3.

(2) Suppose m_1 and m_2 are the same monomial occurrence m in s. If the applied rules are the same then $t_1 = t_2$. So assume that $l_1 \to r_1$ and $l_2 \to r_2$ are different polynomial rewrite rules and let $\langle c_1, c_2\rangle$ be the corresponding critical pair. We have $c_1 \leftarrow lcm(l_1, l_2) \to c_2$ and hence $m'c_1 \leftarrow m \to m'c_2$ where m' is the monomial such that $m = m' \cdot lcm(l_1, l_2)$. By assumption $c_1 \downarrow_\sim c_2$. Clearly $m'c_1 \downarrow_\sim m'c_2$. Since $\to$ and $\sim$ are closed under contexts, we obtain $t_1 = m'c_1 + u \downarrow_\sim m'c_2 + u = t_2$. Here u is the polynomial defined by the equation $s = m + u$.

□

Buchberger's algorithm: inefficient version

Input: • a PRS F
Output: • a Gröbner basis G with the same conversion as F

$C := CP(F)$;
$G := F$;
while $C \neq \varnothing$ **do**
    choose a pair $\langle s, t\rangle \in C$;
    $C := C - \{\langle s, t\rangle\}$;
    reduce $s - t$ to normal form n with respect to G;
    if $\tilde{n} \neq 0$ **then**
        $C := C \cup CP(G, \{\tilde{n}\})$;
        $G := G \cup \{\tilde{n}\}$
    fi
od

FIGURE 11.

Figure 11 shows a simple completion procedure based on the new representation of polynomials. The correctness proof is a straightforward adaptation of the proof of Theorem 4.10.

The reader is invited to check that the developments presented in Sections 5 and 6 can also be performed in the new setting.

References

1. B. Buchberger, *Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory*, in: Multidimensional Systems Theory (ed. N.K. Bose), Reidel, pp. 184-232, 1985.

2. R. Bündgen, *Term Completion versus Algebraic Completion*, Ph.D. thesis, Eberhard-Karls-Universität zu Tübingen, 1991.

3. N. Dershowitz and Z. Manna, *Proving Termination with Multiset Orderings*, Communications of the ACM **22**(8), pp. 465–476, 1979.

4. L.E. Dickson, *Finiteness of the Odd Perfect and Primitive Abundant Numbers with n Distinct Prime Factors*, American Journal of Mathematics **35**, pp. 413–426, 1913.

5. G. Huet, *Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems*, Journal of the ACM **27**(4), pp. 797–821, 1980.

6. D.E. Knuth and P. Bendix, *Simple Word Problems in Universal Algebras*, in: Computational Problems in Abstract Algebra (ed. J. Leech), Pergamon Press, pp. 263–297, 1970.

7. Y. Métivier, *About the Rewriting Systems Produced by the Knuth-Bendix Completion Algorithm*, Information Processing Letters **16**, pp. 31–34, 1983.

8. M.H.A. Newman, *On Theories with a Combinatorial Definition of Equivalence*, Annals of Mathematics **43**(2), pp. 223–243, 1942.

9. F. Winkler and B. Buchberger, *A Criterion for Eliminating Unnecessary Reductions in the Knuth-Bendix Algorithm*, Proceedings of the Colloquium on Algebra, Combinatorics and Logic in Computer Science, Györ, Hungary, 1985.