# Canonical Reduction Systems in Symbolic Mathematics

Franz Winkler

RISC, Johannes Kepler University, Linz, Austria
Franz.Winkler@risc.jku.at
http://risc.uni-linz.ac.at/people/winkler

**Abstract.** Many algorithmic methods in mathematics can be seen as constructing canonical reduction systems for deciding membership problems. Important examples are the Gauss elimination method for linear systems, Euclid's algorithm for computing greatest common divisors, Buchberger's algorithm for constructing Gröbner bases, or the Knuth-Bendix procedure for equational theories. We explain the basic concept of a canonical reduction system and investigate the close connections between these algorithms.

## 1 Introduction

The biological theory of evolution exhibits many instances of similar solutions having been developed for similar problem; examples are the wings of insects, birds, and bats, or the different realizations of light sensitive organs such as eyes. The same phenomenon can be observed in the development of the sciences, and also in particular in mathematics. Many algorithmic methods in different fields of mathematics, e.g. linear algebra, commutative algebra, or logic, can be seen as constructing canonical reduction systems for deciding membership problems. Important examples are the Gauss elimination method for linear systems, Euclid's algorithm for computing greatest common divisors, Buchberger's algorithm for constructing Gröbner bases, or the Knuth-Bendix procedure for equational theories. Here we continue the work started in [4], where we have demonstrated the relations between Buchberger's algorithm for the construction of Gröbner bases and the Knuth-Bendix procedure for the construction of canonical term rewriting systems. We explain the basic concept of canonical reduction systems and investigate the close connections between these algorithms.

### 1.1 Canonical Reduction Relations and Systems

Canonical reduction systems (see also [5], Chapter 8) are supposed to solve the following kind of problem:

- given a mathematical structure $\mathcal{S}$ and a congruence relation $\cong$ on $\mathcal{S}$ (i.e. $\cong \subseteq \mathcal{S}^2$) defined by a finite set of generators $G = \{(l_i, r_i) \mid l_i \cong r_i \text{ for } 1 \leq i \leq n\}$ (i.e. $\cong = \cong_G$),

S. Bozapalidis and G. Rahonis (Eds.): CAI 2009, LNCS 5725, pp. 123–135, 2009.

- we want to construct a new set of generators $\hat{G}$ for this congruence relation, which makes it easy to decide, for any given $s, t \in \mathcal{S}$, whether $s \cong_G t$.

In order to solve such decision problems we introduce a reduction relation $\longrightarrow_G \subseteq \mathcal{S} \times \mathcal{S}$. So, to start with, $\longrightarrow_G$ is simply a binary relation on $\mathcal{S}$. But we will want this relation to have certain properties. Let us first introduce the following notation: for any binary relation $\longrightarrow$ we denote

by $\longleftarrow$ the inverse,
by $\longrightarrow^*$ the reflexive transitive closure, and
by $\longleftrightarrow^*$ the reflexive symmetric transitive closure

of $\longrightarrow$. We will want $\longrightarrow_G$ to have the following properties:

- $\cong_G = \longleftrightarrow_G^*$ , i.e. the symmetric reflexive transitive closure of $\longrightarrow_G$ is equal to the congruence generated by $G$, and
- $\longrightarrow_G$ is **terminating** or **Noetherian**, i.e. every reduction chain $s_0 \longrightarrow_G s_1 \longrightarrow_G \cdots$ is finite.

In addition to being Noetherian, the reduction relation $\longrightarrow_G$ might also be **Church-Rosser**, i.e. $s \longleftrightarrow_G^* t$ implies the existence of a common successor $u$ s.t. $s \longrightarrow_G^* u \longleftarrow_G^* t$. In particular this means that two irreducible elements $s, t$ are congruent if and only if they are syntactically equal.

In case $\longrightarrow_G$ is both Noetherian and Church-Rosser, we call $\longrightarrow_G$ a **canonical reduction relation** and we call $G$ a **canonical reduction system** for the congruence $\cong$.

A canonical reduction system yields the following decision procedure for the underlying congruence $\cong = \longleftrightarrow_G^*$: in order to decide whether $s \cong t$ for $s, t \in \mathcal{S}$,

- reduce $s$ and $t$ to (any) irreducible $s'$ and $t'$ s.t.

$$\begin{aligned} s = s_0 &\longrightarrow_G s_1 \longrightarrow_G \cdots \longrightarrow_G s_m = s', \\ t = t_0 &\longrightarrow_G t_1 \longrightarrow_G \cdots \longrightarrow_G t_n = t' \end{aligned}$$

  ($s'$ and $t'$ are called **normal forms** of $s$ and $t$, respectively);
- check whether $s' = t'$; if so then $s \cong_G t$, otherwise not.

### 1.2 Generating Canonical Reduction Systems

In general a given set of generators $G$ (or its corresponding reduction relation $\longrightarrow_G$) for a congruence $\cong$ will not have the Church-Rosser property. So our goal now becomes to transform $G$ into an equivalent canonical system $\hat{G}$. It turns out that the Church-Rosser property is equivalent to the simpler property of **confluence**, meaning that if $s, t$ have a common predecessor in finitely many steps, $s \uparrow_G^* t$, then they also have a common successor, $s \downarrow_G^* t$. Furthermore, under the assumption of Noetherianity, confluence is equivalent to **local confluence**, meaning that if $s, t$ have a common predecessor in a single step, $s \uparrow_G t$, then they also have a common successor, $s \downarrow_G^* t$. In many interesting cases, such as the

algorithms discussed in this paper, the test for local confluence can be reduced to the test of finitely many **critical pairs**. These are pairs $(s,t)$ s.t. $s \uparrow_G t$, and all other such situations can be regarded as specializations of critical pairs. So if we can prove that for all critical pairs there are common successors, then we have a canonical reduction system. Otherwise we take normal forms $s' \neq t'$ of $s,t$, respectively, and add the pair $(s',t')$ to $G$. Since obviously $s' \longleftrightarrow^*_G t'$, we also have $s' \cong t'$; so the addition of this new pair to $G$ will not violate the requirement $\cong = \longleftrightarrow^*_G$. Of course we have to ensure that by this modification of the reduction system $G$ the Noetherianity of $\longrightarrow_G$ is preserved. In general this is hard, indeed undecidable in some cases such as term rewriting systems (cf. [1],[2]). In any case, we keep considering critical pairs, adding new pairs to the set of generators $G$, and in this way creating more critical pairs. So the question is whether this process will ever terminate and deliver a canonical reduction system. Indeed, for the cases of Gauss elimination, Euclid's algorithm, and Gröbner bases, such a canonical system will finally be produced. But in the case of the Knuth-Bendix procedure for term rewriting systems, the completion process might not yield such a canonical system in finitely many steps.

Let us now demonstrate this approach for the cases listed above.

## 2 Gauss Elimination in Linear Algebra

We consider the following setting:

- the mathematical structure $\mathcal{S}$ is a finite dimensional vector space $V$ over a field $K$; w.l.o.g. $V = K^n$;
- as the generating elements for the congruence we take a basis $B$ for a subvectorspace $W = \text{span}(B)$;
- now the equivalence relation is $v_1 \cong_W v_2 \iff v_1 - v_2 \in W$; and $\cong_W$ is generated by $b \cong_W 0$ for $b \in B$.

The central problem then is to decide whether, for given $v \in V$,

$$v \cong_W 0 \text{ , i.e. } v \in \text{span}(B) = W \ .$$

Every basis $B$ of $W$ generates this congruence; simply let $v$ be congruent w.r.t. $B$ to $w$ if $v - w$ is a linear combination of $B$. We write $\cong_B$ for this congruence, and we observe that $\cong_B = \cong_W$ for every basis $B$ of $W$.

If the basis $B$ (considered as lines of a matrix) is triangular, then this central problem becomes easily decidable. The triangulation or elimination method of Gauss transforms $B$ into such a triangular basis. Let us see that what Gauss elimination does is exactly the construction of a canonical reduction system.

The basis $B$ induces a reduction relation $\longrightarrow_B$ on $V$ as follows:

- for a non-zero vector $b = (0,\ldots,0,b_i,\ldots,b_n)$ with $b_i \neq 0$ we say $\text{lead}(b) = i$;
- now the reduction relation $\longrightarrow_b$ by a single vector $b$ is

$$c = (c_1,\ldots,c_i \neq 0,\ldots,c_n) \ \longrightarrow_b \ c - \frac{c_i}{b_i} \cdot b$$

and for a finite set $B$ we say

$$c \longrightarrow_B d \quad \Longleftrightarrow \quad \exists b \in B : c \longrightarrow_b d \ .$$

It is not hard to see that for every $B$ the reduction relation $\longrightarrow_B$ has the following properties:

- $\longrightarrow_B$ is terminating
- $v \longleftrightarrow_B^* w$ if and only if $v \cong_B w$.

But $\longrightarrow_B$ in general is **not** confluent. Consider the following example: let

$$B = \{\underbrace{(1,0,0)}_{b_1}, \underbrace{(1,1,1)}_{b_2}\}$$

be a basis for a subvectorspace $W = \mathrm{span}(B)$ of $\mathbb{Q}^3$. Then

$$w_1 = (0,2,2) \longleftarrow_{b_1} (1,2,2) = v \longrightarrow_{b_2} (0,1,1) = w_2$$

and both reduction results are irreducible. So $w_1$ and $w_2$ are congruent, $w_1 \cong_B w_2$, but this cannot be determined by reduction w.r.t $B$.

So what can we do in order to transform $\longrightarrow_B$ into a confluent reduction relation? Well, according to Gauss elimination, we consider the elements of $B$ as lines in a matrix (also denoted by $B$) and transform the matrix

$$B \ = \ \begin{pmatrix} b_1 \\ \cdots \\ b_m \end{pmatrix}$$

to row echelon form. This means we look at situations, where the component of a vector can be reduced by (at least) two different generators $b_j$ and $b_k$. Clearly we can simplify this situation to a situation of critical pairs, where a unit vector

$$e_i = (0, \ldots, 0, \underbrace{1}_{i\text{–th position}}, 0, \ldots, 0) \ ,$$

can be reduced by two different generators $b_j$ and $b_k$. This means that $\mathrm{lead}(b_j) = i = \mathrm{lead}(b_k)$, and

$$e_i - b_j \longleftarrow_{b_j} e_i \longrightarrow_{b_k} e_i - b_k$$

(here we have assumed, w.l.o.g., that the components of both $b_j$ and $b_k$ at their leading positions are 1). These reduction results are congruent w.r.t. $\cong_B$, so their difference $b_{m+1} := b_j - b_k$ is in $W$. If $b_{m+1} = 0$, then there was no divergence anyway; otherwise we add $b_{m+1}$ to the basis $B$, thereby enforcing this particular divergence of reduction to converge:

$$\begin{array}{ll} \text{either} & e_i - b_j \longrightarrow_{b_{m+1}} e_i - b_k \\ \text{or} & e_i - b_k \longrightarrow_{b_{m+1}} e_i - b_j \end{array}$$

Observe that this represents exactly a step in the formation of the row echelon form of the matrix (basis) $B$.

This process terminates and yields a set of generators $\hat{B}$ s.t.

- $\longleftrightarrow_B^* \;=\; \cong_W \;=\; \longleftrightarrow_{\hat{B}}^*$ ,
- $\longrightarrow_{\hat{B}}$ is both Noetherian and confluent.

So we can decide the membership problem for $W$ by reduction w.r.t. $\hat{B}$.

For our example above this means the following:

$$\begin{array}{rl} B = \{b_1, b_2\}: & b_1 = \quad (1,0,0) \\ & b_2 = \quad (1,1,1) \\ & - - - \; - - - - - \\ & b_3 = \quad (0,1,1) \\ & \qquad\qquad\qquad \rightarrow \hat{B} = \{b_1, b_2, b_3\} \end{array}$$

Now $\hat{B}$ spans the same vector space $W$, and we can use the reduction w.r.t.$\hat{B}$ to decide membership in $W$:

$$\begin{array}{l} (1,2,2) \longrightarrow_{b_1} (0,2,2) \longrightarrow_{b_3} (0,0,0) \\ (1,2,2) \longrightarrow_{b_2} (0,1,1) \longrightarrow_{b_3} (0,0,0) \end{array}$$

The vector $(1,2,2)$ is indeed in $W$.

In the end we can clean up the basis by keeping only one element with the same lead; in a confluent system we would never need the others, because in every reduction in which we might want to use one of these basis elements, we might instead use the one we keep. Such an interreduced basis $\hat{B}$ is basically the Hermite matrix associated to $B$.

## 3 Euclid's Algorithm for gcds of Univariate Polynomials

We consider the following setting:

- the mathematical structure $\mathcal{S}$ is $K[x]$, the ring of polynomials over a field $K$;
- as the generating elements for the congruence we take two (or finitely many) non-zero polynomials $F = \{f_1(x), f_2(x)\} \subset K[x]$, generating an ideal $I = \langle F \rangle$ in $K[x]$; $F$ is called a basis for the ideal $I$;
- now the equivalence relation is $h_1 \equiv_I h_2 \iff h_1 - h_2 \in I$ .

The central problem then is to decide whether, for given $h \in K[x]$,

$$h \equiv_I 0 \text{ , i.e. } h \in \langle F \rangle = I \ .$$

If $g$ is the greatest common divisor (gcd) of $f_1$ and $f_2$, then $\langle g \rangle = I = \langle f_1, f_2 \rangle$, and the central question can be easily decided as

$$h \equiv_I 0 \iff g | h \ .$$

The Euclidean algorithm computes exactly this gcd, by a sequence of remainders. W.l.o.g. assume that the degree of $f_1$ is at least as high as the degree of $f_2$. We

let $r_1 := f_1, r_2 := f_2$ be the first two remainders in our sequence; an $r_{i+2}$ is then simply the remainder of $r_i$ on division by $r_{i+1}$. Throughout the algorithm we always have

$$\gcd(f_1, f_2) \;=\; \gcd(r_i, r_{i+1}) \; .$$

It is easy to see that this process of remaindering must terminate with, say, $r_k \neq 0$, but $r_{k+1} = 0$. Then we have

$$\gcd(f_1, f_2) \;=\; \gcd(r_k, 0) \;=\; r_k \; .$$

Throughout the Euclidean algorithm the ideal $I$ remains unchanged, since all these remainders clearly are in $I$.

An ideal basis $F$ induces a reduction relation $\longrightarrow_F$ on $K[x]$ as follows:

- for a non-zero polynomial $f(x) = f_n x^n + \cdots f_1 x + f_0$ with $f_n \neq 0$ we say $\mathrm{lead}(f) = \deg(f) = n$;
- now the reduction relation $\longrightarrow_f$ by a single polynomial $f$ is

$$p = p_m x^m + \cdots + \underbrace{p_i}_{\neq 0} x^i + \cdots + p_0 \;\longrightarrow_f\; p \;-\; \frac{p_i}{f_n} x^{i-n} f(x), \qquad \text{if } i \geq n$$

and for a finite basis $F$ we say

$$p \longrightarrow_F q \quad \Longleftrightarrow \quad \exists f \in F : p \longrightarrow_f q \; .$$

It is not hard to see that for every $F$ the reduction relation $\longrightarrow_F$ has the following properties:

- $\longrightarrow_F$ is terminating, and
- $p \longleftrightarrow_F^* q$ if and only if $p \equiv_I q$.

But $\longrightarrow_F$ in general is **not** confluent. Consider the following example: let

$$F = \{\underbrace{x^5 + x^4 + x^3 - x^2 - x - 1}_{f_1}, \; \underbrace{x^4 + x^2 + 1}_{f_2}\}$$

be a polynomial basis for the ideal $I = \langle f_1, f_2 \rangle$. Then

$$\underbrace{-x^3 + x^2 + x + 2}_{q_1} \longleftarrow_{f_2} -x^4 - x^3 + x + 1 \longleftarrow_{f_1} \underbrace{x^5 - x^2}_{p} \longrightarrow_{f_2} \underbrace{-x^3 - x^2 - x}_{q_2}$$

and both reduction results are irreducible. So $q_1$ and $q_2$ are congruent, $q_1 \equiv_I q_2$, but this cannot be determined by reduction w.r.t. $F$.

So what can we do in order to transform $\longrightarrow_F$ into a confluent reduction relation? Well, we consider terms of least degree which can be reduced by two different polynomials. All other diverging reductions can be seen as derived from such divergences. W.l.o.g. we may assume that all polynomials in our remainder

sequence are monic; instead of the actual remainder, we simply take its monic associate. If $d_i = \deg(f_i)$, then $x^{d_i}$ can be reduced both by $f_i$ and $f_{i+1}$:

$$x^{d_i} - f_i \longleftarrow_F x^{d_i} \longrightarrow_F x^{d_i} - x^{d_i - d_{i-1}} \cdot f_{i+1} \ .$$

If these reduction results are the same, then this divergence of reduction converges, and we are done. Otherwise we add the difference $f_i - x^{d_i - d_{i+1}} f_{i+i}$ to the basis; this obviously leaves the ideal $I$ unchanged. In fact we might as well reduce both sides to normal forms, and then add their difference to the basis. What we have done is simply a step in the division algorithm. In this way we consider all pairs of polynomials in the basis (it can be demonstrated that considering subsequent remainders is sufficient); i.e. we compute a remainder sequence starting with $f_1, f_2$:

$$\begin{array}{lll} F = \{f_1, f_2\} : f_1 & & \\ \quad f_2 & & \\ \quad - - - & & \\ \quad f_3 & := \mathrm{rem}(f_1, f_2) & \\ \quad \vdots & & \\ \quad f_k & (\neq 0) & \\ \quad f_{k+1} & (= 0) & \hat{F} = \{f_1, f_2, \ldots, f_k\} \end{array}$$

This process terminates and yields a set of generators $\hat{F}$ containing $f_k = \gcd(f_1, f_2)$. In fact we have

- $\longleftrightarrow_F^* \ = \ \equiv_I \ = \ \longleftrightarrow_{\hat{F}}^*$ , and
- $\longrightarrow_{\hat{F}}$ is both Noetherian and confluent.

So we can decide the membership problem for $I$ by reduction w.r.t. $\hat{F}$:

$$h \in \langle F \rangle \iff f_k | h \iff h \longrightarrow_{\hat{F}} 0 \ .$$

For our example above this means the following:

$$\begin{array}{lll} F = \{f_1, f_2\} : f_1 = & x^5 + x^4 + x^3 - x^2 - x - 1 & \\ \quad f_2 = & x^4 + x^2 + 1 & \\ \quad - - - - - - - - & & \\ \quad f_3 = & x^4 - x^2 - 2x - 1 = & f_1 - x \cdot f_2 \\ \quad f_4 = & x^2 + x + 1 = & \frac{1}{2}(f_2 - f_3) \\ \quad f_5 = & 0 = & f_3 - (x^2 - x - 1) f_4 \\ & & \to \hat{F} = \{f_1, \ldots, f_4\} \end{array}$$

Now $\hat{F}$ generates the same ideal $I$, and we can use the reduction w.r.t. $\hat{F}$ to decide membership in $I$:

$$0 \longleftarrow_{f_3} -x^3 + x^2 + x + 2 \longleftarrow_{f_1, f_2} x^5 - x^2 \longrightarrow_{f_2} -x^3 - x^2 - x \longrightarrow_{f_3} 0 \ .$$

So $x^5 - x^2 \in I$.

In the end we can again interreduce the elements in the confluent reduction system $\hat{F}$. Whenever we might want to use a basis polynomial different from $f_k$ in a reduction, we might as well use $f_k$. So since our reduction system is now confluent, we don't need the other basis polynomials any more; we simply keep $\hat{F} = \{f_k\}$.

## 4 Gröbner Bases in Multivariate Polynomial Rings

We consider the following setting:

- the mathematical structure $\mathcal{S}$ is $K[x_1, \ldots, x_n]$, the ring of multivariate polynomials over a field $K$;
- as the generating elements for the congruence we take finitely many non-zero polynomials $F = \{f_1, \ldots, f_m\} \subset K[x_1, \ldots, x_n]$ generating an ideal $I = \langle F \rangle$ in $K[x_1, \ldots, x_n]$;
- now the equivalence relation is $h_1 \equiv_I h_2 \iff h_1 - h_2 \in I$ .

The central problem then is to decide whether, for given $h \in K[x_1, \ldots, x_n]$,

$$h \equiv_I 0 \text{ , i.e. } h \in \langle F \rangle = I \ .$$

As in the case of univariate polynomials we would like to introduce a reduction w.r.t. a basis $F$, and then add certain polynomials to the basis in order to make the corresponding reduction relation confluent. Such an ideal basis we will then call a **Gröbner basis**. Buchberger's algorithm for the construction of Gröbner bases does exactly that.

For introducing a reduction relation $\longrightarrow_F$, we first have to linearly order the multivariate terms $x_1^{e_1} \cdots x_n^{e_n}$. In the univariate case we did not have any choice; the only reasonable ordering is induced by the degree. But in the multivariate case we have much more freedom. We need to choose an ordering respecting the multiplicative structure of the set of terms, called an **admissible ordering**; i.e.

- $1 = x^{(0,\ldots,0)} \leq s$ for every term $s$, and
- if $s \leq t$ and $u$ any term, then $s \cdot u \leq t \cdot u$.

There is an abundance of such admissible orderings; e.g. lexicographic orderings, graduated lexicographic orderings, and many others. Admissible orderings are completely classified. Once we have chosen an admissible ordering $<$ of the terms, every non-zero polynomial $f$ has a well-defined **leading term** $\mathrm{lead}(f)$ (the highest term in the ordering appearing with non-zero coefficient in $f$) and a non-zero **leading coefficient** $\mathrm{lc}(f)$, the coefficient of $\mathrm{lead}(f)$. By $\mathrm{le}(f)$ we denote the exponent (vector) of $\mathrm{lead}(f)$.

Now we are ready for defining the reduction relation $\longrightarrow_F$ on $K[x_1, \ldots, x_n]$ (for the fixed admissible term ordering $<$): for a non-zero polynomial

$$p = p_{\mathrm{le}(p)} x^{\mathrm{le}(p)} + \cdots + p_e x^{e=(e_1,\ldots,e_n)} + \cdots \ , \qquad \text{with } p_e \neq 0$$

we define

$$p \longrightarrow_f p - \frac{p_e}{\mathrm{lc}(f)} x^{e-\mathrm{le}(f)} f(x), \qquad \text{if } e - \mathrm{le}(f) \in \mathbb{N}^n$$

and $$p \longrightarrow_F q \quad \Longleftrightarrow \quad \exists f \in F : p \longrightarrow_f q\,.$$

Again, as in the univariate case, one can prove that $\longrightarrow_F$ has the following properties:

- $\longrightarrow_F$ is terminating, and
- $p \longrightarrow_F^* q$ if and only if $p \equiv_I q$.

But $\longrightarrow_F$ in general is **not** confluent. Consider the following example: let

$$F = \{\underbrace{x^2y^2 + y - 1}_{f_1},\ \underbrace{x^2y + x}_{f_2}\}$$

be a basis for the polynomial ideal $I = \langle f_1, f_2\rangle$. Then

$$q_1 = -y + 1 \longleftarrow_{f_1} p = x^2y^2 \longrightarrow_{f_2} -xy = q_2$$

and both results are irreducible. So $q_1$ and $q_2$ are congruent, $q_1 \equiv_F q_2$, but this cannot be determined by reduction w.r.t. $F$.

So what do we do in order to transform $\longrightarrow_F$ into a confluent reduction relation? Well, as in the previous cases (Gauss elimination, Euclidean algorithm) we investigate the "smallest" situations in which something can be reduced in essentially two different ways. We look at terms $x^e$ which can be reduced w.r.t. two different generators $f_j, f_k$. This means that $\mathrm{lead}(f_j)|x^e$ and also $\mathrm{lead}(f_k)|x^e$. The (finitely many) smallest such situations occur when $x^e = \mathrm{lcm}(\mathrm{lead}(f_j), \mathrm{lead}(f_k))$ (least common multiple), and all the other cases are instantiations of such basic situations (see [5] for details). We reduce $x^e$ both modulo $f_j$ and $f_k$, getting some $g_j$ and $g_k$, respectively. $g_j$ and $g_k$ may be further reduced modulo $\longrightarrow_F$ to normal forms $g_j'$ and $g_k'$, respectively:

$$g_j' \longleftarrow_F^* g_j \longleftarrow_{f_j} x^e = \mathrm{lcm}(\mathrm{lead}(f_j), \mathrm{lead}(f_k)) \longrightarrow_{f_k} g_k \longrightarrow_F^* g_k'$$

Actually we reduce $g_j - g_k$, the so-called **S-polynomial** of $f_j$ and $f_k$, to a normal form $h$. If $h = 0$, then this divergence of reduction converges, and we are done. Otherwise we observe that $h \in I$. So if we add $h$ to the basis $F$, then this divergence can be resolved, and the ideal remains unchanged.

Of course, now we have a new element in the basis, and there are more S-polynomials to be considered. But this process terminates and yields a set of generators $\hat{F}$ s.t.

- $\longleftrightarrow_F^* \;=\; \equiv_I \;=\; \longleftrightarrow_{\hat{F}}^*$, and
- $\longrightarrow_{\hat{F}}$ is both Noetherian and confluent.

So we have computed a Gröbner basis $\hat{F}$ for the ideal $I$ w.r.t. the term ordering $<$. With the Gröbner basis $\hat{F}$ for $I$, we can decide the membership problem for $I$ by reduction w.r.t. $\hat{F}$. If in the end we interreduce the elements in $\hat{F}$, we get a **minimal Gröbner basis** for the ideal $I$.

For our example above this means the following. We choose an admissible term ordering, say graduated lexicographic with $x < y$. Then we consider S-polynomials and reduce them to normal forms. This leads to the following sequence of polynomials being added to the basis:

$$\begin{array}{lllll} F: & f_1 = & x^2y^2 + y - 1 & & \\ & f_2 = & x^2y + x & & \\ & - - - & - - - - - & & \\ & f_3 = & -xy + y - 1 = & f_1 - y \cdot f_2 & \\ & f_4 = & y - 1 = & f_2 + (x+1)f_3 & \\ & f_5 = & -x = & f_3 + (x-1)f_4 & \\ & & & \to \hat{F} = \{f_1, \ldots, f_5\} & \end{array}$$

Now $\hat{F}$ generates the same ideal $I$, and we can use the reduction w.r.t. $\hat{F}$ to decide membership in $I$:

$$\begin{array}{llll} x^2y^2 & \longrightarrow_{f_1} -y + 1 & \longrightarrow_{f_4} 0 \\ x^2y^2 & \longrightarrow_{f_2} -xy & \longrightarrow_{f_5} 0 \end{array}$$

So $x^2y^2 \in I$. The minimal Gröbner basis for $I$ is $\{x, y - 1\}$.

## 5 The Knuth-Bendix Procedure for Term Rewriting Systems

We consider the following setting:

- a term algebra $\mathcal{T}(\Sigma, V)$ over a signature $\Sigma$ and variables $V$;
- $E = \{s_i = t_i \mid i \in I\}$ a set of equations over $\mathcal{T}$ generating an equational theory $=_E$ ;
- now the equivalence relation is $s \equiv_E t \iff s = t \in =_E$ .

The equational theory $=_E$ is the set of all equations which can be derived from $E$ by reflexivity, symmetry, transitivity, substitution, and replacing equals by equals; confer [1], [2].

The central problem then is to decide whether, for given $s, t \in \mathcal{T}(\Sigma, V)$,

$$s =_E t \ .$$

We define a reduction relation on $\mathcal{T}(\Sigma, V)$ by orienting the equations in $E$

$$e_i: \quad s_i = t_i$$

in one of the ways (according to a reduction ordering)

$$r_i: \quad s_i \longrightarrow t_i \quad \text{or} \quad t_i \longrightarrow s_i$$

(w.l.o.g. assume $r_i : s_i \longrightarrow t_i$). This leads to a so-called **rewrite rule system (RRS)**

$$R = \{r_i \mid i \in I\} \ .$$

The reduction $\longrightarrow_R$ works in the following way: if there is a substitution $\sigma$ and a position $p$ in the term $u$, such that $\sigma$ applied to $s_i$ equals the subterm of $u$ at position $p$, i.e. $\sigma(s_i) = u_{|p}$, then this subterm of $u$ can be replaced by $\sigma(t_i)$:

$$u \longrightarrow_R v \quad \Longleftrightarrow \quad \exists p, i, \sigma : u_{|p} = \sigma(s_i), \text{ and } v = u[p \leftarrow \sigma(t_i)] \ .$$

Here $u[p \leftarrow \sigma(t_i)]$ means that in $u$ we replace the subterm at position $p$ by the term $\sigma(t_i)$.

In general the termination property is undecidabel for rewrite rule systems. But there are several sufficient conditions; e.g. $s_i > t_i$ w.r.t. a reduction ordering. For the following let us assume that the rules can be ordered w.r.t. such a reduction ordering. Then $\longrightarrow_R$ has the following properties:

- $\longrightarrow_R$ is terminating, and
- $\longleftrightarrow_R^* = =_E$ .

But $\longrightarrow_R$ in general is **not** confluent. Consider the example of group theory; i.e. let $G$ consist of the axioms

$$\begin{array}{rcl} G & = & \{\ (1)\ \ 1 \cdot x = x, \\ & & \ \ (2)\ \ x^{-1} \cdot x = 1, \\ & & \ \ (3)\ \ (x \cdot y) \cdot z = x \cdot (y \cdot z)\ \ \} \ , \end{array}$$

which are oriented (lexicographic path ordering with $^{-1} > \cdot > 1$) to give the rewrite rule system

$$\begin{array}{rcl} R & = & \{\ (1)\ \ 1 \cdot x \longrightarrow x, \\ & & \ \ (2)\ \ x^{-1} \cdot x \longrightarrow 1, \\ & & \ \ (3)\ \ (x \cdot y) \cdot z \longrightarrow x \cdot (y \cdot z)\ \ \} \ . \end{array}$$

Then

$$x^{-1} \cdot (x \cdot y) \ \longleftarrow_{(3)} \ (x^{-1} \cdot x) \cdot y \ \longrightarrow_{(2)} \ 1 \cdot y \ \longrightarrow_{(1)} \ y$$

Both results are irreducible, they are congruent modulo $=_E$, but they have no common successor.

So again the goal is to transform the RRS $R$ into an equivalent confluent RRS $\hat{R}$,

$$\longleftrightarrow_R^* \ = \ \longleftrightarrow_{\hat{R}}^* \ .$$

As in the previous cases (Gauss elimination, Euclidean algorithm, Gröbner bases) we investigate "smallest" situations in which a term can be reduced in essentially two different ways. Towards this end, we consider (not necessarily different) rules

$$r : s \longrightarrow t \ , \qquad r' : s' \longrightarrow t' \ ,$$

a most general unifier (substitution) $\sigma$, and a position $p$ in the term $s$ ($s_{|p}$ not being a variable) s.t.

$$\sigma(s') \ = \ \sigma(s_{|p}) \ .$$

In this case we get the following divergence in reduction

$$v = \sigma(t) \longleftarrow_r \sigma(s) = u \longrightarrow_{r'} \sigma(s[p \leftarrow t']) = v' \ .$$

The pair of terms $(v, v')$ is called a **critical pair** of the RRS $R$. The components of the critical pair $(v, v')$ are obviously equal modulo $=_E$; so are normal forms $w$ and $w'$ to which $v$ and $v'$ can be reduced, respectively. If $w \neq w'$, then we try to orient them into a new rule $w \longrightarrow w'$ or $w' \longrightarrow w$, which does not violate the termination property of the RRS.

In contrast to the previous cases (Gauss elimination, Euclidean algorithm, Gröbner bases), there is no guarantee that this completion process will terminate. Critical pairs will lead to new rules, which lead to new critical pairs, which will lead to new rules, and so on. Also we might get stuck in a situation where the normal forms of a critical pairs, $w$ and $w'$, cannot be oriented into a rule without violating the termination property. But if this process terminates and yields a RRS $\hat{R}$ then

- $\longleftrightarrow^*_R \ = \ =_E \ = \ \longleftrightarrow^*_{\hat{R}}$, and
- $\longrightarrow_{\hat{R}}$ is both Noetherian and confluent.

So we can decide the equality modulo $E$ by reduction w.r.t. $\hat{R}$. In the end we can interreduce the RRS $\hat{R}$ and so get a minimal RRS for $=_E$.

For the example of group theory this means that because of

$$x^{-1} \cdot (x \cdot y) \ \longleftarrow_{(3)} \ (x^{-1} \cdot x) \cdot y \ \longrightarrow_{(2)} \ 1 \cdot y \ \longrightarrow_{(1)} \ y$$

we add the new rule

$$(4) \ x^{-1} \cdot (x \cdot y) \ \longrightarrow \ y \ .$$

We continue to consider other critical pairs. For the case of group theory this completion process (according to Knuth and Bendix, cf. [3]) actually terminates and yields the following minimal rewrite rule system:

$$\begin{array}{ll}
(1) & 1 \cdot x \longrightarrow x, \\
(2) & x^{-1} \cdot x \longrightarrow 1, \\
(3) & (x \cdot y) \cdot z \longrightarrow x \cdot (y \cdot z), \\
(4) & x^{-1} \cdot (x \cdot y) \longrightarrow y, \\
(5) & x \cdot 1 \longrightarrow x, \\
(6) & 1^{-1} \longrightarrow 1, \\
(7) & (x^{-1})^{-1} \longrightarrow x, \\
(8) & x \cdot x^{-1} \longrightarrow 1, \\
(9) & x \cdot (x^{-1} \cdot y) \longrightarrow y, \\
(10) & (x \cdot y)^{-1} \longrightarrow y^{-1} \cdot x^{-1}.
\end{array}$$

So the equational theory of pure group theory can be decided by reduction modulo this RRS. Also for many other equationally definable algebraic structures there are canonical rewrite rule systems.

## 6 Conclusion

We have seen that several key algorithms in constructive algebra and logic actually are based on the same idea; namely the formation of critical pairs and the completion of a reduction relation. Recognition of these similarities might lead to a better understanding of algorithms and perhaps to new application areas. And mathematics can be seen as a more unified and interrelated field of knowledge.

**Acknowledgement.** This work has been supported by the Austrian Science Fund (FWF) under the project DIFFOP, No. P20336-N18.

## References

1. Avenhaus, J.: Reduktionssysteme. Springer, Heidelberg (1995)
2. Book, R.V., Otto, R.: String-Rewriting Systems. Springer, Heidelberg (1993)
3. Knuth, D.E., Bendix, P.B.: Simple Word Problems in Universal Algebra. In: Leech, J. (ed.) Computational Problems in Abstract Algebra, pp. 263–297. Pergamon Press, Oxford (1970)
4. Winkler, F.: The Church–Rosser property in computer algebra and special theorem proving: An investigation of critical–pair/completion algorithms, Dissertation Univ. Linz, Austria, VWGÖ Wien (1984)
5. Winkler, F.: Polynomial Algorithms in Computer Algebra. Springer, Wien New York (1996)