

G22.3033-003 Logic and Verification

Spring 2004

Lecture 11

1

Outline

- Gröbner bases

Sources:

Harrison, John. *Introduction to Logic and Automated Theorem Proving*.
Unpublished manuscript. Used by permission.

Gallian, Joseph A. *Contemporary Abstract Algebra, Second Edition*.
Heath and Company, 1990.

3

Review

- Algebraically Closed Fields

2

Gröbner bases

Last time, we gave a general quantifier elimination method for algebraically closed fields.

However, each quantifier eliminated often makes the formula larger and increases the degree of the other variables.

If there are a significant number of quantifiers, this tends to make the procedure very inefficient.

A natural question is whether there might be a more efficient method, say for eliminating more than one quantifier at once.

In fact, there are more efficient general methods for eliminating a whole block of nested existential quantifiers.

We will consider a special case that is quite straightforward: *eliminating a block of existential quantifiers in a formula containing no other variables*.

Notice, that in particular, this will give an improved decision procedure for universal formulas.

4

Gröbner bases

Consider eliminating a block of quantifiers from a formula of the form

$$\exists x_1 \dots, x_n. \bigwedge_i p_i(x_1, \dots, x_n) = 0 \wedge \bigwedge_j q_j(x_1, \dots, x_n) \neq 0.$$

We first note that we can eliminate all disequalities using the following transformation:

$$q(x_1, \dots, x_n) \neq 0 \leftrightarrow \exists z. q(x_1, \dots, x_n) \cdot z + 1 = 0.$$

This is called the *Rabinowitsch trick* and is widely used as a theoretical device.

Note that if we cannot eliminate a block of existentials all at once, this trick is of dubious usefulness, since it adds an existential quantifier.

However, if we can eliminate blocks of quantifiers, this technique is likely preferable to the approach of multiplying all the disequalities together discussed last time.

This is because the degrees of the polynomials are increased by only one in this method, whereas they may be increased significantly by multiplying polynomials together.

5

Gröbner bases

Polynomial Rings

Recall that a ring is any structure that satisfies the following axioms.

- $\forall x y z. x + (y + z) = (x + y) + z$
- $\forall x. 0 + x = x$
- $\forall x. (-x) + x = 0$
- $\forall x y. x + y = y + x$
- $\forall x y z. x \times (y \times z) = (x \times y) \times z$
- $\forall x y z. x \times (y + z) = x \times y + x \times z$

Note that given an appropriate canonizer, the set of polynomials in \vec{x} whose coefficients are from a ring is itself a ring.

Thus, given a ring R , we can talk about the ring $R[\vec{x}]$ of polynomials in the variables \vec{x} .

7

Gröbner bases

Let $\vec{x} = x_1, \dots, x_n$. We are left with the problem of eliminating the existential quantifiers from

$$\exists \vec{x}. p_1(\vec{x}) = 0 \wedge \dots \wedge p_k(\vec{x}) = 0.$$

where we have made the assumption that the equations contain no other variables besides those in \vec{x} .

In other words, we must find whether a set of polynomial equations has a common solution.

This is a fundamental problem in algebraic geometry, and the following theorem gives an equivalent condition.

(Weak) Hilbert Nullstellensatz (Zero-point Theorem)

The polynomial equations $p_1(\vec{x}) = 0, \dots, p_k(\vec{x}) = 0$ in an algebraically closed field have no common solution iff there are polynomials $q_1(\vec{x}), \dots, q_k(\vec{x})$ such that the following identity holds:

$$q_1(\vec{x}) \cdot p_1(\vec{x}) + \dots + q_k(\vec{x}) \cdot p_k(\vec{x}) = 1.$$

6

Gröbner bases

Ideals

Given a ring R , an *ideal* of R is a nonempty subset A of R such that $0 \in A$, A is closed under $+$, and for every $r \in R$ and $a \in A$, $r \cdot a \in A$.

Given $a_1, \dots, a_k \in R$, the ideal $\langle \{a_1, \dots, a_k\} \rangle$ generated by $\{a_1, \dots, a_k\}$ is the set of all $r_1 a_1 + \dots + r_k a_k$, where $r_1, \dots, r_k \in R$.

Thus, suppose we have $p_1(\vec{x}), \dots, p_k(\vec{x})$, polynomials over an algebraically closed field F . These are also elements of the polynomial ring $F[\vec{x}]$.

Then, the set of all $q_1(\vec{x}) \cdot p_1(\vec{x}) + \dots + q_k(\vec{x}) \cdot p_k(\vec{x})$ is exactly the ideal of $F[\vec{x}]$ generated by $\{p_1(\vec{x}), \dots, p_k(\vec{x})\}$.

The question of whether there exist polynomials $q_1(\vec{x}), \dots, q_k(\vec{x})$ such that

$$q_1(\vec{x}) \cdot p_1(\vec{x}) + \dots + q_k(\vec{x}) \cdot p_k(\vec{x}) = 1,$$

is exactly the question of whether 1 is in the ideal of $F[\vec{x}]$ generated by $\{p_1(\vec{x}), \dots, p_k(\vec{x})\}$.

8

Gröbner bases

The problem of determining ideal membership for multivariate polynomials over a field was solved in 1965 by **Buchberger** in his PhD thesis (Gröbner was his PhD advisor).

The basic idea of Buchberger's algorithm is to use polynomials as rewrite rules to reduce other polynomials.

The goal is to obtain a canonical reduction system:

A set R of polynomials is a *Gröbner basis* for an ideal J if $J = \langle R \rangle$ (i.e. J is generated by R) and R defines a canonical (terminating and confluent) rewrite system.

The methods used by Buchberger are related to those used in Knuth-Bendix completion, though Buchberger's algorithm predated the Knuth-Bendix completion work.

Because of the similarities, we will build on the rewriting framework described earlier to explain Buchberger's algorithm.

We will first show how to obtain a Gröbner basis from an arbitrary set R of polynomials. We will then show how this solves the ideal membership problem.

9

Gröbner bases

Now we can see how to use canonical polynomial equations as rewrite rules.

Given an equation $m_1 + m_2 + \dots + m_p = 0$, we can consider it as an equation $m_1 = -m_2 + \dots + -m_p$.

Now, any polynomial which contains any multiple of m_1 can be rewritten:

$$qm_1 \longrightarrow -qm_2 + \dots + -qm_p.$$

Example

As an example, consider rewriting $x^4 + 1$ with $x^2 = xy - y$:

$$\begin{aligned}
x^4 + 1 &\longrightarrow x^2(xy - y) + 1 \\
&= x^3y - x^2y + 1 \\
&\longrightarrow xy(xy - y) - x^2y + 1 \\
&= x^2y^2 - x^2y - xy^2 + 1 \\
&\longrightarrow y^2(xy - y) - x^2y - xy^2 + 1 \\
&= xy^3 - x^2y - xy^2 - y^3 + 1 \\
&\longrightarrow xy^3 - y(xy - y) - xy^2 - y^3 + 1 \\
&= xy^3 - 2xy^2 - y^3 + y^2 + 1
\end{aligned}$$

11-g

Gröbner bases

Canonical Forms

As with previous procedures, it will be convenient to define a canonical form for polynomial terms.

The form we choose is a finite sum of *monomials*, each of which is of the form $ax_1^{m_1} \dots x_k^{m_k}$, i.e. some element a of the field multiplied by some powers (possibly 0) of each of the variables.

A canonical form for polynomial equations is $m_1 + \dots + m_p = 0$ where each m_i is a monomial and all monomials whose powers of the variables are equivalent have been combined.

There remains the question of what order to list the monomials in. For reasons which will become clear later, we desire a total ordering with the property that if $m_1 < m_2$, then for any other monomial m , we have $m \cdot m_1 < m \cdot m_2$. This property is called *compatibility*.

An obvious ordering satisfying the compatibility criterion is to compare monomials first according to the *multidegree*, the sum of the degrees of all the variables, and then lexicographically according to some fixed order of the variables.

10

Gröbner bases

Given a set of polynomial equations $\bigwedge_i p_i(\vec{x}) = 0$, we can turn each one into a rewrite rule to obtain a set R of rewrites.

Recall that by Newman's lemma, if \longrightarrow_R is terminating and weakly confluent, then it is canonical: $R \models s = t$ iff $s \longleftarrow_R^* t$ iff $s \downarrow_R t$.

The multidegree ordering we defined is well-founded, so \longrightarrow_R is automatically terminating. To determine weak confluence, we can consider a finite number of critical situations, similar to the critical pairs of Knuth-Bendix completion.

Suppose p, q_1, q_2 are polynomials and R is a set of polynomial rewrite rules, and suppose that $p \longrightarrow_R q_1$ and $p \longrightarrow_R q_2$. We would like to figure out if q_1 and q_2 are joinable. There are two possibilities:

- The reductions result from rewriting different monomials, i.e. $p = m_1 + m_2 + p_0$ and one rewrite maps $m_1 \longrightarrow_R r_1$ and the other maps $m_2 \longrightarrow_R r_2$. Thus, $q_1 = r_1 + m_2 + p_0$ and $q_2 = m_1 + r_2 + p_0$.
- The reductions result from rewriting the same monomial, i.e. $p = m + p_0$ and one rewrite maps $m \longrightarrow_R r_1$ and the other maps $m \longrightarrow_R r_2$. Thus, $q_1 = r_1 + p_0$ and $q_2 = r_2 + p_0$.

12

Gröbner bases

Consider the first case:

$p = m_1 + m_2 + p_0$, where $m_1 \rightarrow_R r_1$ and $m_2 \rightarrow_R r_2$, so

$q_1 = r_1 + m_2 + p_0$ and $q_2 = m_1 + r_2 + p_0$.

One might suspect that q_1 and q_2 are always joinable simply because the other monomial can still be rewritten.

However, notice that once q_1 and q_2 are put in canonical form, the other monomial might disappear (if, for example $-m_2$ appears in r_1).

It turns out that such instances are always joinable, but for more subtle reasons.

Suppose $m_1 > m_2$. Now notice that r_2 cannot contain m_1 because of the ordering constraint.

It is possible, however, that r_1 contains m_2 , so that $r_1 = am_2 + s_2$ for some rational a and some other polynomial s_2 .

13

Gröbner bases

Thus non-confluence can only occur in the second situation, in which rewrites are applied to the same monomial.

To handle these cases, we have to look at the *S-polynomials* of the rewrite system. These are roughly analogous to critical pairs in equational logic.

Given two polynomials p and q which define rewrites $m_1 \rightarrow p_1$ and $m_2 \rightarrow p_2$, define their *S-polynomial* as

$$S(p, q) = p_1 m'_1 - p_2 m'_2,$$

where $LCM(m_1, m_2) = m_1 m'_1 = m_2 m'_2$.

Theorem

Suppose R is a set of polynomial equations and that for any two polynomials $p, q \in R$, $S(p, q) \rightarrow_R^* 0$. Then \rightarrow_R is confluent.

Before proving this, we need a number of preliminary results.

Note that the converse is also true, but we will not prove it.

15

Gröbner bases

We have:

$q_1 = r_1 + m_2 + p_0$, $q_2 = m_1 + r_2 + p_0$, and $r_1 = am_2 + s_2$ where $m_1 > m_2$, $m_1 \rightarrow_R r_1$ and $m_2 \rightarrow_R r_2$.

Then,

$$\begin{aligned} q_1 &= r_1 + m_2 + p_0 \\ &= (am_2 + s_2) + m_2 + p_0 \\ &= (a+1)m_2 + s_2 + p_0 \\ &\rightarrow_R^* (a+1)r_2 + s_2 + p_0. \end{aligned}$$

Similarly,

$$\begin{aligned} q_2 &= m_1 + r_2 + p_0 \\ &\rightarrow_R r_1 + r_2 + p_0 \\ &= am_2 + s_2 + r_2 + p_0 \\ &\rightarrow_R^* ar_2 + s_2 + r_2 + p_0 \\ &= (a+1)r_2 + s_2 + p_0. \end{aligned}$$

14-h

Gröbner bases

Lemma

If $p \rightarrow_R q$ and m is a nonzero monomial, then also $mp \rightarrow_R mq$.

Corollary

If $p \rightarrow_R^* q$ and m is a monomial or zero, then also $mp \rightarrow_R^* mq$.

Theorem

If $p - q \rightarrow_R^* 0$, then $p \downarrow_R q$.

Proof

The proof is by induction on the length of the reduction sequence in $p - q \rightarrow_R^* 0$.

For the base case, if $p - q = 0$, then $p = q$ and the result is trivial.

Otherwise, suppose $p - q \rightarrow_R r \rightarrow_R^* 0$. The step $p - q \rightarrow_R r$ must arise from some rewrite $m \rightarrow_R s$ where some multiple of m appears in $p - q$ and thus in at least one of p and q .

16

Gröbner bases

We have:

$$p - q \rightarrow_R r \rightarrow_R^* 0$$

$$m \rightarrow_R s$$

$$p = am + p_1$$

$$q = bm + q_1$$

$$p - q = (a - b)m + (p_1 - q_1)$$

$$r = (a - b)s + (p_1 - q_1)$$

Let $p' = as + p_1$ and $q' = bs + q_1$ and notice that using the same rewrite $m \rightarrow_R s$, we get $p \rightarrow_R^* p'$ and $q \rightarrow_R^* q'$.

But $p' - q' = r$, so we know that $p' - q' \rightarrow_R^* 0$. It follows from the induction hypothesis that that $p' \downarrow_R q'$.

But since $p \rightarrow_R^* p'$ and $q \rightarrow_R^* q'$, this shows that $p \downarrow_R q$. \square

17

Gröbner bases

Thus, we can test whether R is a Gröbner basis by checking all its S-polynomials.

But we can do better. As with critical pairs, we can use S-polynomials to complete R : if reduction of $S(p, q)$ leads to an irreducible polynomial $r \neq 0$, we simply add r to R , with the result that $S(p, q) \rightarrow_R^* 0$.

Notice that by adding r to R , we do not change the ideal $\langle R \rangle$. This is because r is already in $\langle R \rangle$. To verify this, we first show $S(p, q) \in \langle R \rangle$.

Recall that $p = m_1 - p_1$, $q = m_2 - p_2$, and $S(p, q) = p_1m'_1 - p_2m'_2$, where $LCM(m_1, m_2) = m_1m'_1 = m_2m'_2$.

$$\begin{aligned} S(p, q) &= p_1m'_1 - p_2m'_2 \\ &= p_1m'_1 - m_1m'_1 - p_2m'_2 + m_1m'_1 \\ &= p_1m'_1 - m_1m'_1 - p_2m'_2 + m_2m'_2 \\ &= -m'_1(p) + m'_2(q). \end{aligned}$$

Now, notice that every application of a rewrite is equivalent to adding some multiple of a polynomial in R , i.e. if $s \in \langle R \rangle$ and $s \rightarrow_R s'$, we have $s' = s + mt$, for some monomial m and $t \in R$. Thus $s' \in \langle R \rangle$.

It follows that $r \in \langle R \rangle$.

19-d

Gröbner bases

We can now prove the main theorem:

Suppose R is a set of polynomial equations and that for any two polynomials $p, q \in R$, $S(p, q) \rightarrow_R^* 0$. Then \rightarrow_R is confluent.

Proof

Recall that the only possibility for nonconfluence is when two polynomials q_1 and q_2 cause rewrites $m_1 \rightarrow q'_1$ and $m_2 \rightarrow q'_2$ which apply to the same monomial m in a polynomial $p = m + p'$.

Since m must be a multiple of both m_1 and m_2 , it must also be a multiple of $LCM(m_1, m_2)$, so we can write $m = m' LCM(m_1, m_2)$.

Letting $LCM(m_1, m_2) = m_1m'_1 = m_2m'_2$, we have $p \rightarrow_R m'q'_1m'_1 + p'$ and $p \rightarrow_R m'q'_2m'_2 + p'$.

Now, by hypothesis, we have $S(q_1, q_2) \rightarrow_R^* 0$, so $q'_1m'_1 - q'_2m'_2 \rightarrow_R^* 0$. It follows by the above corollary that $m'q'_1m'_1 - m'q'_2m'_2 \rightarrow_R^* 0$, and thus $(m'q'_1m'_1 + p') - (m'q'_2m'_2 + p') \rightarrow_R^* 0$.

But, by the above theorem, this implies that $m'q'_1m'_1 + p' \downarrow_R m'q'_2m'_2 + p'$, and thus \rightarrow_R is confluent. \square

18

Gröbner bases

Finally, we show that in contrast to Knuth-Bendix completion, we can always achieve confluence by adding a finite number of polynomials derived from S-polynomials.

Each polynomial r added to R is irreducible with respect to R . In particular, it has no monomial divisible by the head monomial of any existing polynomial in R .

Thus, nontermination of the algorithm implies the existence of an infinite sequence of monomials m_i such that $m_i \not\prec m_j$ for $i < j$.

A monomial $cx_1^{m_1} \cdots x_k^{m_k}$ divides a monomial $dx_1^{n_1} \cdots x_k^{n_k}$ iff $m_i \leq n_i$ for all $1 \leq i \leq k$.

Termination is then an immediate consequence of the following result.

Dickson's Lemma

Define the ordering \leq_n on \mathcal{N}^n by $(x_1, \dots, x_n) \leq (y_1, \dots, y_n)$ iff $x_i \leq y_i$ for all $1 \leq i \leq n$. Then there is no infinite sequence t_i of elements of \mathcal{N}^n such that $t_i \not\leq t_j$ for all $i < j$.

We omit the proof which is by induction on n .

20

Gröbner bases

Thus, given any set R of polynomials, we can compute a Gröbner basis for $\langle R \rangle$. This solves the problem of determining membership of a polynomial in $\langle R \rangle$:

Theorem

If R is a Gröbner basis for $\langle R \rangle$ and p is any polynomial, then $p \rightarrow_R^* 0$ iff $p \in \langle R \rangle$.

Before proving this theorem, we remind ourselves why we care.

The original problem was to eliminate a block of existential quantifiers from $\exists \vec{x}. \phi(\vec{x})$, where we assume that ϕ contains no variables besides those in \vec{x} .

Using the Rabinowitsch trick, we can transform this into the problem of determining $\exists \vec{x}. p_1(\vec{x}) = 0 \wedge \dots \wedge p_k(\vec{x}) = 0$.

The weak zero point theorem tells us that this is false iff there are polynomials $q_1(\vec{x}), \dots, q_k(\vec{x})$ such that $q_1(\vec{x}) \cdot p_1(\vec{x}) + \dots + q_k(\vec{x}) \cdot p_k(\vec{x}) = 1$.

But this is true iff 1 is in the ideal $\langle p_1(\vec{x}) \dots p_k(\vec{x}) \rangle$.

By the above theorem, we can determine this by computing the Gröbner basis R for $\langle p_1(\vec{x}) \dots p_k(\vec{x}) \rangle$ and then checking whether $1 \rightarrow_R^* 0$.

Gröbner bases

The following two lemmas can be proved by induction.

Lemma

If \rightarrow_R is confluent, and $p \downarrow_R q$, then for any polynomial r , $p + r \downarrow_R q + r$.

Lemma

If \rightarrow_R is confluent and $p \downarrow_R q$, then $rp \downarrow_R rq$ for any polynomial r .

Theorem

If R is a Gröbner basis for $\langle R \rangle$ and p is any polynomial, then $p \rightarrow_R^* 0$ iff $p \in \langle R \rangle$.

Proof

To see the 'only if' part, first note that if $p \rightarrow_R q$, then $q = p + mr$ for some monomial m and $r \in R$. So $p - q = -mr \in \langle R \rangle$. It follows that if $p \rightarrow_R^* q$, then $p - q \in \langle R \rangle$. Thus, if $p \rightarrow_R^* 0$, then $p \in \langle R \rangle$.

In the other direction, if $p \in \langle R \rangle$ then $p = \sum_{i=1}^k q_i p_i$ where each $p_i \in R$. Since trivially each $p_i \rightarrow_R 0$, it follows by the above lemmas that $p \rightarrow_R^* 0$. \square