Overview of Quantifier Elimination

George E. Collins

September 15, 1998

An *atomic formula* is any of the formulas $A > 0$, $A = 0$, $A < 0$, $A \geq 0$, $A \leq$ and $A \neq 0$, where A is any polynomial in several variables with integer coefficients. A *quantifier-free formula* of the elementary theory of the real numbers is any formula constructed from atomic formulas using the Boolean connectives such as $\wedge, \vee, \sim, \rightarrow$ and $\equiv$. A *formula* of this theory is obtained by quantifying a quantifier-free formula, applying existential and universal quantifiers to any of the variables occurring in a quantifier-free formula. A *statement* of the theory is a formula in which all variables are bound by quantifiers. Although the coefficients of the polynomials are integers, the variables and the quantifiers are regarded as ranging over the field of all real numbers. A *quantifier elimination method* for the theory is an algorithm which takes as input any formula of the theory and produces as output an equivalent quantifier-free formula containing only the free (unquantified) variables of the input formula. A sentence is either true or false, so a quantifier elimination method is a *decision method.* The output formula corresoponding to a sentence contains no variables, which is easily reducible to either $0 = 0$ (true) or $0 \neq 0$ (false). The first quantifier elimination method was discovered by Alfred Tarski, published in 1951 /citeTar51. In 1973 the author discovered a new quantifier elimination method, [1] with greater promise of practical application.

Since 1975 many significant improvements have been made to this method, and it has been implemented as an application program, **qepcad**, of SACLIB. As one application of the program, consider the following instance of the so-called implicitization problem. We wish to determine the set of all real points (x, y, z) such that

$$(\exists u)(\exists v)[x = 3u + 3uv^2 - u^3 \wedge y = 3v + 3u^2v - v^3 \wedge z = 3u^2 - 3v^2].$$

After 32.7 seconds of computation, the program produces the answer

$$64z^9 - 10368z^7 + 1296y^2z^6 - 1296x^2z^6 - 34992y^2z^5$$

$$-34992x^2z^5 + 419904z^5 + 174960y^2z^4 - 174960x^2z^4$$

$$-10935y^4z^3 - 56862x^2y^2z^3 + 314928y^2z^3 - 10935x^4z^3$$

$$+314928x^2z^3 + 118098y^4z^2 - 118098x^4z^2 + 59049y^4z$$
$$-118098x^2y^2z + 59049x^4z + 19683y^6 - 59049x^2y^4$$
$$+59049x^4y^2 - 19683x^6 = 0.$$

Collins' method of quantifier elimination is called cylindrical algebraic decomposition, abbreviated CAD. CAD is an algorithm is an algorithm that given as input a number of variables, say n, an ordered list of n variables, say $(x_1, \ldots, x_n)$, and a list of polynomials, say $(A_1, \ldots, A_k)$, produces a decomposition of n-dimensional real space into a finite number of connected sets, in each of which each input polynomial has invariant sign. Further, these sets have a cylindrical structure determined by the input list of variables. The sign-invariance and the cylindrical structure crucial to the quantifier elimination, and the ordering of the variables must be consistent with the order in which the quantifiers are applied in the input formula. To illustrate the idea, let $n = 2$, $k = 1$, let (x, y) be the variable list, and let $A_1(x, y)$ be the tacnode polynomial

$$y^4 - 2y^3 + y^2 - 3x^2y + 2x^4.$$

qepcad produces the decomposition shown in Figure 1 in 100 milliseconds. This figure consists of 11 cylinders. There are 5 vertical lines, each of which is a cylinder, and they divide the plane into the 4 cylindrical regions between the adjacent vertical lines, one cylindrical region to the left of the leftmost line, and one to the right of the rightmost line. Each cylinder contains some of the sign-invariant connected regions, which are called *cells*. The leftmost cylinder consists of a single cell, in which the tacnode polynomial is positive. The next cylinder, a line, consists of 3 cells, the middle cell consisting of a single point where the tacnode is zero. In the two cells on either side the tacnode is positive. In the third cylinder there are five cells, two of them being portions of the tacnode curve where the polynomial is zero. Between these two cells the tacnode is negative, and in the two remaining cells the tacnode is positive. Now suppose that the input formula to **qepcad** is

$$(\exists y)[y^4 - 2y^3 + y^2 - 3x^2y + 2x^4 < 0].$$

The program will determine that the values of x satisfying the formula are precisely those in all cylinders except the two leftmost, the two rightmost, and the center line (the y-axis), just by observing which cylinders contain at least one cell where the tacnode polynomial is negative. This information can be used to produce a formula in x alone that is satisfied by precisely those values of x. That formula will contain the polynomial $2048x^6 - 4608x^4 + 37x^2 + 12$ and the polynomial x. These are the irreducible factors of the discriminant with respect to y of the tacnode polynomial. The real roots of these two polynomials are the values of x in the vertical lines of Figure 1.

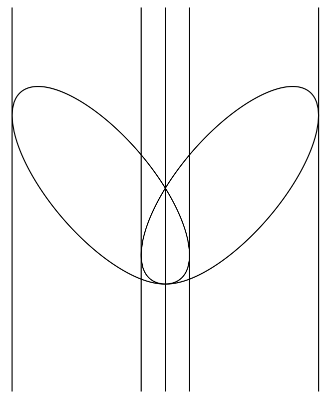

Figure 1: The CAD of the tacnode.

Cylindrical algebraic decomposition is very useful apart from its application to quantifier elimination. It can be used, for example, to determine the number of connected components of any curve or surface, or to plan robot motions. These applications require, in addition, an algorithm for deciding whether any two cells of a CAD are adjacent in the sense that their union is a connected set.

Cylindrical algebraic decomposition and quantifier elimination require many algebraic algorithms. These include computation of polynomial resultants and discriminants, complete factorization of polynomials, computation of polynomial gcd's (greatest common divisors), isolation of the real roots of polynomials with either integer or real algebraic number coefficients, and computation of primitive elements for multiple algebraic extension fields. All of these algorithms will be studied in the coming lectures. These algorithms must also be extremely efficient because the polynomials arising in the CAD process are enormous for interesting applications. Therefore the study of the computational requirements of these algorithms will also be given much attention.

References

[1] George E. Collins Quantifier Elimination for Real Closed Fields by Cylindrical Algebraic Decomposition. In: Brakhage, H. (Ed.), Automata Theory and Formal Languages. Springer, pp. 134-183 (Lecture Notes in Computer Science, vol. 33). Reprinted in Quantifier Elimination and Cylindrical Decomposition, Springer, 1998.

[2] A. Tarski. A decision method for elementary algebra and geometry. University of California Press, 1951. Reprinted in Quantifier Elimination and

Cylindrical Decomposition, Springer, 1998.