Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT®

Journal of Complexity 19 (2003) 631–637

Journal of COMPLEXITY

http://www.elsevier.com/locate/jco

Use of algebraically independent numbers for zero recognition of polynomial terms

Daniel Richardson*,[1] and Ahmed El-Sonbaty[2]

Department of Computer Science, Bath University, Bath BA2 7AY, UK

Abstract

A polynomial term is a tree with operators $*, +, -$ on the interior nodes and natural numbers and variables on the frontier.

We attempt to decide whether or not such a tree represents the zero polynomial by substituting algebraically independent real numbers for the variables and attempting to decide whether or not the resulting constant is zero.

From this we get a probabilistic zero recognition test which is somewhat more expensive computationally than the usual method of choosing random integers in a large interval and evaluating, but which depends on the ability to choose a random point in the unit cube and to approximate a polynomial at that point.

We also state a conjecture about algebraic independence measure which would give us a deterministic test with a uniformly chosen test point. The result is that if a polynomial term has $s(T)$ nodes, then the bit complexity of deterministic zero recognition is bounded by $O(s(T)M(s(T)\,length(T)))$, where $length(T)$ measures the length of the term T, and $M(n)$ is the bit complexity of multiplication of two n-digit natural numbers.

1. Introduction

Consider representation of polynomials as trees with $+, -, *$ on the interior nodes and with variables and natural numbers on the leaves. Such a representation is significantly more succinct than any of the usual canonical representations of

*Corresponding author. Department of Computer Science, Bath University, Bath BA2 7AY, UK.

E-mail addresses: masdr@bath.ac.uk (D. Richardson), mapaaeee@bath.ac.uk (A. El-Sonbaty).

[1] Supported by EC Research Training Network HPRN-CT-2001-00271 RAAG.

[2] Supported by the Egyptian Government.

0885-064X/03/$ - see front matter © 2003 Elsevier Inc. All rights reserved.

doi:10.1016/S0885-064X(03)00047-5

$\mathbf{Z}[x_1, \ldots, x_n]$. On the other hand, it is much weaker than the standard straight line program representation.

A probabilistic method which can be used for zero recognition among such terms is independently to pick n integers $(x_1, \ldots, x_n)$ at random in the range $[0, d * n * N]$, where d is an upper bound on the degree and N is some large multiplier, and then substitute these integers into the polynomial and evaluate. The following theorem implies that the probability of an error (caused by accidentally choosing a root of the polynomial) would be no more than $(1/N)$, provided that the polynomial is independent of the random number generator.

Let $deg(T, x_i)$ be the degree of variable x_i in polynomial term T. Suppose T has variables $x_1, \ldots, x_n$. Define $d(T) = \sum_{i=1}^{n} deg(T, x_i)$.

Theorem 1. *Let $p \in A[x_1, \ldots, x_n]$ be a not identically zero polynomial over an integral domain A. Let S be a subset of A of cardinality B. Suppose $x_1, \ldots, x_n$ are chosen independently, with uniform probability, from S. Then the probability that $p(x_1, \ldots, x_n) = 0$ is bounded by $d(T)/B$.*

See [Zippel] for a proof.

Define $h(n)$ where n is a natural number to be the number of digits used in the canonical base 10 representation of n. We will call $h(n)$ the logarithmic height of n. For polynomial terms T, define $h(T)$ to be the maximum of $h(n)$ for n on the frontier of T. Define $length(n) = h(n)$, and $length(x_n) = h(n) + 1$. Extend this notion of length to polynomial terms by setting $length(A + B) = length(A - B) = length(A * B) = length(A) + length(B) + 1$.

According to the above theorem, if we want the probability of an error to be, for example, less than 10^{-100}, we need to evaluate the polynomial with n natural number arguments of logarithmic height no more than $100 + h(d(T))$. After substitution, the length of the term is bounded by $length(T) * (50 + h(d(T)))$. Assume that there are no more than $s(T)$ nodes in the tree. The largest integer in the computation has length no more than $length(T) * (50 + h(d(T)))$. There are no more than $s(T)$ computations to be done. So the total complexity is bounded by $s(T)M(length(T) * (50 + h(d(T))))$, where $M(k)$ is the complexity of multiplication of two natural numbers of length k.

We next prove an analogous theorem about points X chosen at random in the unit cube in R^n.

Let $\{|T(X)| < 10^{-k}\}$ be the subset of R^n in which $|T(X)| < 10^{-k}$.

Theorem 2. *If polynomial term T is not identically zero, then the intersection of the unit cube in R^n with $\{|T(X)| < 10^{-k}\}$ has Lebesgue measure $\leqslant 2d(T)10^{-k/d(T)}$.*

Proof. We use induction on n and the fact that if $|a_1 \ldots a_d| > |b_1 \ldots b_d|$ then for some i we must have $|a_i| > |b_i|$.

Assume $T(X)$ is not identically zero. Let x be one of the variables in T, $d = deg(T, x)$, $D = d(T) - d$.

For almost all values of the variables,

$$T(X) = a_d(x-\alpha_1)\dots(x-\alpha_d),$$

where $\alpha_1, \dots, \alpha_d$ depend algebraically on the other variables, and a_d is a nonzero polynomial in the other variables.

$$\begin{aligned}&\{|T(X)|<10^{-k}\}\\ &\quad=\{|a_d||m(x)|<10^{-k}\}\\ &\quad\subseteq\{|a_d|<10^{-k\alpha}\}\cup\{|m(x)|<10^{-k\beta}\},\end{aligned}$$

where $\alpha = D/(d+D)$, $\beta = d/(d+D)$, and $m(x) = (x-\alpha_1)\dots(x-\alpha_d)$.

Observe that $\{m(x)<10^{-k\beta}\}\subseteq\bigcup_{i=1}^{d}\{|x-\alpha_i|<10^{-k\beta/d}\}$.

Let L be Lebesgue measure, and I^n the unit cube in R^n.

We apply our induction hypothesis to get an upper bound on $L(I^n\cap\{|a_d|<10^{-k\alpha}\})$ (using the fact that the measure of a projection parallel to the x-axis of a subset of the unit cube is an upper bound on the measure of the subset) and then apply the observation above.

$$\begin{aligned}&L(I^n\cap\{|T(X)|<10^{-k}\})\\ &\quad\leqslant L(I^n\cap\{|a_d|<10^{-k\alpha}\})+L(I^n\cap|m(x)|<10^{-k\beta}\})\\ &\quad\leqslant 2D10^{-k\alpha/D}+2d10^{-k\beta/d}\\ &\quad\leqslant 2(d+D)10^{-k/(d+D)}\end{aligned}$$

(since α/D and β/d are both $1/(d+D)$), which verifies the theorem. □

Remark. For the sake of simplicity the theorem is stated in terms of the unit cube. We could change the scale so that it would apply to any cube $[0,B]^n$ in R^n by multiplying the measure by a factor of B^n.

To compare with the method stated earlier, suppose we choose X at random and we want the probability of an error to be less than 10^{-100}. According to the theorem it would be acceptable, from this point of view, to assume that $T=0$ if $|T(X)|<10^{-k}$, where $k/d(T)>h(2d(T))+100$. Therefore it suffices to choose:

$$k>d(T)(h(2d(T))+100).$$

We have a bound of $B=10^{length(T)}$ on the absolute values of the numbers which appear in the tree.

At each operation $+,*,-$ within the polynomial term we lose no more than $length(T)$ decimal places of precision.

During the whole computation the total precision lost is bounded by $s(T)*length(T)$.

It suffices, according to the above theorem, to finish the computation with precision k.

The precision needed for the whole computation is bounded by

$$precision = s(T) * length(T) + d(T) * (h(2d(T)) + 100).$$

Since we have $s(T)$ interior nodes (operations), the bit complexity of the computation is $s(T)\, M(precision)$. It can be seen that this test is more expensive computationally than the previously described integer test.

2. Choice of points in the unit cube

When we choose integers to test whether or not a polynomial term is zero, we make a different random choice each time we have a term to test. Otherwise, someone could discredit the test by constructing a polynomial which happened to be zero at the test point. However, it may not be necessary to choose a new point for each test if we use a randomly chosen point in the unit cube. That is, it may be possible to use always the same point.

Suppose that $k(T)$, mapping polynomial terms into natural numbers, is such that $\sum d(T)10^{-k(T)}$ (taken over all polynomial terms) converges to a finite limit. (For example, $k(T) = 2 * length(T)$ would do.) It is a consequence of the theorem above that for almost all points X in the unit cube there is a number N so that for all nonzero polynomial terms T, $d(T) > N \to |T(X)| \geqslant 10^{-k(T)d(T)}$.

We also have:

Theorem 3. *For almost all points X in the unit cube, there is a number C so that for all nonzero polynomial terms T we have*

$$|T(X)| > 10^{-C-k(T)d(T)}.$$

Proof. Let $(T_i)_{i>0}$ enumerate all the polynomial terms. We assume that $\sum d(T_i)10^{-k(T_i)}$ converges. This implies that for any $\varepsilon > 0$ there is an N such that

$$\sum_{i>N} d(T_i)10^{-k(T_i)} < \varepsilon.$$

This means, as a consequence of the theorem above, that the points X so that $|T_i(X)| < 10^{-k(T)d(T)}$ for some T_i with $i > N$, have measure no more than ε. We can choose ε as small as we wish. The probability is therefore zero that $|T(X)| < 10^{-k(T)d(T)}$ for infinitely many T, when X is chosen at random. Let S be a subset of the unit cube of measure 1 so that for points X picked in S there are only finitely many polynomial terms T with $|T(X)| < 10^{-k(T)d(T)}$. For each point in S there is a constant C so that $|T(X)| > 10^{-C-k(T)d(T)}$ for all polynomial terms T which are not the zero polynomial. □

Thus if we had a way of computing with any of these very common values, and we could find an appropriate constant C, we could have a deterministic zero test for

polynomial terms, and always use the same test point. This possibility is discussed below.

3. Algebraic independence

Definition 1. Complex numbers $a_1, \ldots, a_n$ are algebraically independent if $p(a_1, \ldots, a_n) \neq 0$ for all not identically zero polynomials p in $\mathbf{Z}[x_1, \ldots, x_n]$.

3.1. Lindemann's theorem

Theorem 4. *For any distinct algebraic numbers $\alpha_1, \ldots, \alpha_n$ and nonzero algebraic numbers $\beta_1, \ldots, \beta_n$ we have*

$$\beta_1 e_1^{\alpha} + \cdots + \beta_n e_n^{\alpha} \neq 0.$$

Corollary 1. *If the algebraic numbers $\alpha_1, \ldots, \alpha_n$ are linearly independent over* $\mathbf{Q}$, *then $e^{\alpha_1}, \ldots, e^{\alpha_n}$ are algebraically independent.*

For proof, see [Baker].

The following is well known. See [Besicovitch].

Lemma 1. *If $q_1, \ldots, q_n$ are different square free natural numbers, then $1/\sqrt{q_1}, \ldots, 1/\sqrt{q_n}$ are linearly independent over the rationals.*

4. A conjecture about independence

We extend our notion of length by defining $length(A^{1/n})$ to be $length(A) + length(n) + 1$, $length(1/A) = length(e^A) = length(A) + 1$, and $length(A + B) = length(A - B) = length(A * B) = length(A) + length(B) + 1$.

Conjecture. *Let $T(x_1, \ldots, x_n)$ be a polynomial term, and let A be the term which is obtained by substituting $e^{-1/\sqrt{q_1}}, \ldots, e^{-1/\sqrt{q_n}}$ for $x_1, \ldots, x_n$ respectively, where $q_1, \ldots, q_n$ are natural numbers. Then*

$$A \neq 0 \;\rightarrow\; |A| > 10^{-2^{\,length(A)}}.$$

This is a special version of the uniformity conjecture, which is part of an attempt to solve the zero recognition problem for some of the constants which appear frequently in scientific computing. See [Richardson97,Richardson2001,Richardson2000,RichardsonLa] for discussion of this problem and the conjecture. This is also related to a family of conjectures, called witness conjectures, stated by Van Der Hoeven, see [VDHoeven]. The general form of the uniformity conjecture, and the

strongest version of the witness conjecture, have recently been shown to be incorrect, via a counterexample found by Van Der Hoeven. However the more specialised conjecture above still seems plausible.

We note that the substitution chosen is in the unit cube.

Theorem 5 (Using the conjecture). *Let T be a polynomial term, and let A be the term obtained by the substitution used in the conjecture above, with $q_1, \ldots, q_n$ the first n square free numbers. Then T is identically zero if and only if*

$$|A| < 10^{-10\ length(T)}.$$

Proof. The substituted values are algebraically independent, as a consequence of the Lindemann theorem and the lemma of Besicovitch stated in the previous section. So T represents the zero polynomial if and only if A is zero. According to the conjecture, A is zero if and only if $|A| < 10^{-2\ length(A)}$. The first n square free numbers all have length bounded by n. (This follows from the Bertrand postulate. See [Zippel]. Much stronger results can be obtained by more careful consideration of the density of the primes.) So $length(e^{-1/\sqrt{q_n}})$ is bounded by $h(n) + 5$, and therefore $length(A)$ is bounded by $5\,length(T)$. □

Corollary 2 (Using the conjecture). *There is a deterministic test for zero equivalence of a polynomial term T which has bit complexity which is polynomial in $length(T)$.*

Proof. We need to approximate A with precision at least $k = 10 * length(T)$. The numbers which occur in the computation have absolute value bounded by $10^{length(T)}$. Therefore the precision lost at each step of the computation is bounded by $length(T)$. There are $s(T)$ steps. Thus the total precision lost in the computation is bounded by $s(T)length(T)$. We need to do the whole computation with precision no more than $k_2 = (s(T) + 10)\ length(T)$. At the beginning of the computation, we approximate the numbers $e^{-1/\sqrt{q_i}}$ with precision k_2. Using the results in [Borwein,Borwein1988,Brent] we can do this in $O(M(k_2))$ bit operations, where $M(k)$ is the bit complexity of multiplying two k-digit natural numbers. Thus the whole computation has bit complexity bounded by $O(s(T)M(s(T)\ length(T)))$, which is bounded by a polynomial in the length of the term T.

Note that a much weaker form of the conjecture will still give a polynomial time deterministic solution of the zero recognition problem for polynomial terms. The bound k on the precision could be any polynomial in the length of T. For example, the bound obtained in the first part of this paper for the random substitution from the unit cube would also establish the corollary. In spite of this, we have not so far been able to establish the existence of such a deterministic decision procedure without use of some as yet unproved statement of independence measure such as the conjecture above. We also note the somewhat surprising fact that the *nonexistence* of a polynomial time deterministic solution to the zero recognition problem for

polynomial terms would have very interesting consequences for independence measure of many familiar numbers, including exponentials of algebraic numbers.

We remark also that our zero recognition algorithm for polynomial terms is in complexity class NC, since there exists an efficient parallelisation of evaluation of such terms. See [Brent74,Brent].

References

[Baker] A. Baker, Transcendental Number Theory, Cambridge University Press, Cambridge, 1975.

[Besicovitch] A.S. Besicovitch, On the linear independence of fractional powers of integers, J. London Math. Soc. 15 (1940) 3–6.

[Borwein] J.M. Borwein, P.B. Borwein, Pi and the AGM, Wiley, Canadian Mathematical Society, 1987.

[Borwein1988] J.M. Borwein, P.B. Borwein, On the complexity of familiar functions and numbers, SIAM Rev. 30 (4) (1988) 589–601.

[Brent74] R.P. Brent, The parallel evaluation of generic arithmetic expressions, J. ACM 21 (1974) 201–206.

[Brent] R.P. Brent, Multiple-precision zero-finding methods and the complexity of elementary functions, in: J.F. Traub (Ed.), Analytic Computational Complexity, Academic Press, New York, 1975, pp. 151–176.

[Richardson97] D. Richardson, How to recognise zero, J. Symbolic Comput. 24 (1997) 627–645.

[Richardson2000] D. Richardson, The uniformity conjecture, Proceedings of Computability and Complexity in Analysis, CCA2000, September 17–19, Swansea, Wales. Also in associated Springer Lecture Notes in Computer Science, Vol. 2064, Springer, Berlin, 2000, pp. 253–272.

[Richardson2001] D. Richardson, Multiplicative independence of algebraic numbers and expressions, J. Pure Appl. Algebra 164 (2001) 231–245.

[RichardsonLa] in: T. Mora (Ed.), Some Observations about Familiar Numbers, ISSAC2002, ACM Press, pp. 214–220.

[VDHoeven] J. Van Der Hoeven, Automatic asymptotics, Ph.D. Thesis, Ecole Polytechnique, 1997.

[Zippel] R. Zippel, Effective Polynomial Computation, Kluwer Academic Publishers, Dordrecht, 1993.