A Lower Bound on the Size of Algebraic Decision Trees for the MAX Problem

Dima Grigoriev *

Marek Karpinski †

Andrew Yao ‡

Abstract

We prove an exponential lower bound on the size of (ternary) algebraic decision trees for the MAX Problem of finding maximum of n real numbers. This complements $n-1$ lower bound (cf. M. O. Rabin [R72]) on the depth of algebraic decision trees for this problem. The method yields also for the first time a lower size bound for a polyhedral decision problem MAX= of testing whether the ith number is the maximum among a list of n real numbers, and gives the first nonlinear size lower bound on algebraic decision trees for the selection problems.

*Dept. of Computer Science, PennState University, University Park, PA 16802. Supported in part by Volkswagen-Stiftung. Email: `dima@cse.psu.edu`.

†Dept. of Computer Science, University of Bonn, 53117 Bonn, and International Computer Science Institute, Berkeley, California. Research partially done while visiting the Dept. of Computer Science at Princeton University. Research supported in part by DFG Grant KA 673/4-1, by the ESPRIT Grants 7097 and EC-US 030, and by the Volkswagen-Stiftung. Email: `marek@cs.bonn.edu`.

‡Dept. of Computer Science, Princeton University, Princeton, NJ 08544. Research supported in part by NSF Grant CCR-9301430. Email: `yao@cs.princeton.edu`.

1 Introduction

The MAX Problem of finding maximum of n real numbers was studied extensively in the literature over the years (cf., e.g., [K73], [R72], [TY93]). M. Rabin [R72] has established the sharp lower bound $n-1$ on the depth of any algebraic decision tree solving this problem (see also [MPR94]). Surprisingly, H. Ting and A. Yao [TY93] have proved recently $O((\log n)^2)$ upper bound on the depth of randomized algebraic decision algorithms for computing the maximum of n distinct real numbers. The problem of a lower bound on the size of algebraic decision trees remained open since Rabin's result [R72]. Note that the following construction excludes the possibility of applying Rabin's method [R72] for the lower bounds on the size of algebraic decision trees (the method yields $n-1$ lower bound on the length of any path from the root to the leaf only (!) for linear decision trees or decision trees with median tests [Y89]). Let us take an algebraic decision tree with the first test being $p = x_1 - x_2^2 - \cdots - x_n^2 - 1 : 0$. Then the positive branch ($p(x) > 0$) ends in a leaf (is of length 1) with the output $x_1 = MAX$. This 'counting' difficulty has led to the situation that the only known lower bound on the size of algebraic decision trees (and this even for trees with quadratic polynomials) for solving the MAX Problem was $n-1$.

The size of an algebraic decision tree characterizes its algebraic tree-program complexity, and in the case of linear and median-tests decision trees it played an important role in establishing lower bounds for various selection problems (cf. [FG79], [Y89]) with MAX being their generic subproblem. As mentioned before the only known size lower bound on algebraic decision trees for the MAX Problem was a linear one.

In this paper we prove for the first time an exponential lower bound on the size of algebraic decision trees solving the MAX Problem. This gives also exponential lower bounds for other selection problems. The method of our proof introduces a new lower bound technique which can be of independent interest.

We consider in the paper standard ternary decision trees (cf. [SY82]) branching according to the signs $>, =, <$. Notice that for binary decision trees studied in [R72] (branching according to the signs $\leq, >$) the upper size bound $n-1$ for the MAX= Problem is obvious, namely the tree successively tests, $X_1 \geq X_2, \quad X_1 \geq X_3, \quad \ldots, \quad X_1 \geq X_{n-1}$. Moreover, this construction works in the same way for any polyhedron given by inequalities $a_1X \geq b_1, \ldots, a_kX \geq b_k$. The corresponding binary tree has size k (and depth k as well, compare this with the lower bound on the depth ($logN$) of [GKV95] where N is the number of all the faces of a polyhedron). Observe also that if one studies the well-known problem of membership to a union of hyperplanes $\bigcup_{1\leq i\leq k} a_iX = b_i$, then there is a ternary (as well as binary) algebraic decision tree of the size k. Namely, the tree tests $(a_1X - b_1)^2$, then, if the sign is $>$, tests $(a_2X - b_2)^2$, etc. Notice that this tree is nonlinear.

Still it remains an open question whether there exists a binary algebraic decision tree with a subexponential size for solving the MAX Problem.

2 Size of Algebraic Decision Trees for the MAX Problem

An *Algebraic Decision Tree* (ADT) T of degree d and dimension n for *MAX Problem* is a ternary tree with inner nodes being query nodes of polynomials of degree at most d and branching according to the sign. The input of the tree is an n-tupel $(x_1, \ldots, x_n) \in \mathbb{R}^n$, and each leaf of the tree is either labeled by a certain index $i \in \{1, \ldots, n\}$ or by the labels "yes"or "no". We say that the decision tree T solves MAX (MAX=) Problem in dimension n if for an arbitrary vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$ the path in T branching according to this vector terminates in a leaf labeled by i ("yes") if and only if $x_i = \max\{x_1, \ldots, x_n\}$ ($x_1 = \max\{x_1, \ldots, x_n\}$, respectively).

By the size (resp. depth) of T we mean the number of its nodes (resp. the maximum length of its paths). It is well-known [R72] that the depth of an ADT for the MAX Problem is at least $n-1$. This bound is optimal as one can easily construct T with the depth $n-1$ and size 2^{n-1}.

We prove the exponential lower bound on the size.

Theorem. Any decision tree of degree d for MAX (MAX=) problem has the size at least $2^{c'(d)n}$ where $c'(d) > 0$ depends only on d.

The proof will be conducted for the special case of the MAX= Problem.

For any $\{1\} \subset I \subset \{1, \ldots, n\}$ consider the following ('wall') set $M_I = \{(x_1, \ldots, x_n) : x_i > x_j \text{ for any } i \in I, j \notin I \text{ and } x_{i_1} = x_{i_2} \text{ for any } i_1, i_2 \in I\}$. Denote the plane $P_I = \{(x_1, \ldots, x_n) : x_{i_1} = x_{i_2} \text{ for any } i_1, i_2 \in I\}$. Then $\dim P_I = n - |I| + 1$ and M_I is an open polyhedron in P_I. Note that $P_{\{1\}} = \mathbb{R}^n$ and all P_I are pairwise distinct. Obviously, the sets M_I are pairwise disjoint and form a partition of the set $x_1 = max\{x_1, \ldots, x_n\}$ with 2^{n-1} elements. Observe that the Euclidian closure $\overline{M_I}$ has a non-empty intersection with M_J if and only if $I \subset J$. Moreover, if $\overline{M_I} \bigcap M_J \neq \emptyset$ then $\overline{M_I} \supset M_J$. Thus, $\{M_I\}$ form a cellular decomposition of the set $x_1 = max\{x_1, \ldots, x_n\}$ and the boundary $\partial M_I = \bigcup\limits_{J \supset I} M_J$ in the plane P_I.

The method of our proof is based on the analysis of a "touching frequency" of the sets computed along the branches of a tree T with the 'wall sets' M_I.

For a branch B of the tree T let the testing polynomials together with their signs along this branch be $f_1 = \cdots = f_k = 0, \quad g_1 > 0, \ldots, g_l > 0$. By $W_B \subset \mathbb{R}^n$ denote the semialgebraic set $\{f_1 = \cdots = f_k = 0, \quad g_1 > 0, \ldots, g_l > 0\}$. We say that W_B touches M_I if $\dim(W_B \bigcap M_I) = \dim M_I = n - |I| + 1$. Observe that if W_B touches M_I then the label of B is "yes". Since for every M_I there exists B

such that W_B touches M_I then the theorem will follow from the following Main Lemma.

Main Lemma. For any branch B of T, W_B can touch at most $2^{c(d)n}$ sets M_I for some $c(d) < 1$ depending only on d.

Remark. $c(d)$ is determined recursively as follows. $c(1) = \frac{1}{2}$, then $c(d) = c(d-1) + \frac{1-c(d-1)}{1+(\log_2(2^d/2^d-1)/d)}$.

We proceed to the proof of the Main Lemma.

Proposition 1. W_B cannot touch M_I, M_J such that $I \underset{\neq}{\supset} J$.

Proof. Assume the contrary. Let $\{j_1, \dots, j_{n-|J|}\} = \{1, \dots, n\} \setminus J$. For any polynomial $f \in \mathbb{R}[X_1, \dots, X_n]$ denote $f^{(J)}(X_1, X_{j_1}, \dots, X_{j_{n-|J|}}) = f|_{X_j = X_1, j \in J} \in \mathbb{R}[X_1, X_{j_1}, \dots, X_{j_{n-|J|}}]$. One could consider $f^{(J)}$ as the restriction of f on the plane P_J with the coordinates $X_1, X_{j_1}, \dots, X_{j_{n-|J|}}$, where $X_1 = X_j$ for each $j \in J$. Then $f_1^{(J)}, \dots, f_k^{(J)}$ vanish identically because these polynomials vanish on the semialgebraic set $W_B \bigcap M_J$ of the full dimension in the plane P_J.

By assumption there exists a point $x \in M_I$ such that $g_1(x) > 0, \dots, g_\ell(x) > 0$. There exists a ball $B_x(r)$ with a radius $r > 0$ centered in x such that $g_1, \dots, g_l$ are positive everywhere on $B_x(r)$. As $x \in M_I \subset \partial \overline{M_J}$ there exists a point $x' \in (B_x(r) \bigcap P_J) \setminus \overline{M_J}$. The decision tree T being applied to x' goes through the branch B, since $x' = (x'_1, \dots, x'_n) \in W_B$. We get a contradiction with that the label of B is "yes" since $\max(x'_1, \dots, x'_n)$ is not x'_1 as $x' \notin \overline{M_J}$. Proposition is proved. $\square$

Remark. In fact we proved a stronger statement. Namely, if W_B touches M_J then $W_B \bigcap M_I = \emptyset$ for any $I \underset{\neq}{\supset} J$.

Proposition 2. If W_B touches M_I then I is a minimal (with respect to the inclusion) among the subsets $\{1\} \subset J \subset \{1, \dots, n\}$ such that $f_1^{(J)}, \dots, f_k^{(J)}$ vanish identically.

Proof. Firstly, as we have seen in the proof of Proposition 1 that $f_1^{(I)}, \ldots, f_k^{(I)}$ vanish identically. Secondly, assume that $J \subsetneq I$ such that $f_1^{(J)}, \ldots, f_k^{(J)}$ vanish identically. As W_B touches M_I, there exists a point $x \in M_I \bigcap W_B$, then $g_1(x) > 0, \ldots, g_l(x) > 0$. Then $g_1, \ldots, g_l$ are positive everywhere in a ball $B_x(r)$ for a suitable $r > 0$. Since $x \in M_I \subset \partial M_J$, the open set $B_x(r) \bigcap M_J$ in P_J is nonempty, and $B_x(r) \bigcap M_J \subset W_B$ by the assumption. Thus, W_B touches M_J and we get a contradiction with the Proposition 1, which proves the proposition. $\square$

The Main Lemma would follow from the Proposition 2 and the following proposition.

Proposition 3. For any polynomials $h_1, \ldots, h_m \in \mathbb{R}[X_1, \ldots, X_n]$ with degrees $\deg(h_i) \leq d$ the number of sets minimal (with respect to the inclusion) among the subsets $\{1\} \subset I \subset \{1, \ldots, n\}$ such that $h_1^{(I)}, \ldots, h_m^{(I)}$ vanish identically, does not exceed $2^{c(d)n}$.

We prove the proposition by induction on d. For $d = 1$, each $h_i = \sum_{1 \leq j \leq n} \alpha_{ij} X_j + \alpha_i, 1 \leq i \leq m$ is a linear polynomial. Let $\{1\} \subset I \subset \{1, \ldots, n\}$ be a minimal set for which $h_1^{(I)}, \ldots, h_m^{(I)}$ vanish identically. Then $\alpha_i = 0$ and I contains all j such that $\alpha_{ij} \neq 0$ and finally $\sum_{1 \leq j \leq n} \alpha_{ij} = 0$. Thus, I consists of $\{1\}$ and all $j \in \{1, \ldots, n\}$ for which there exists $1 \leq i \leq m$ such that $\alpha_{ij} \neq 0$, provided that $\alpha_i = 0, \sum_{1 \leq j \leq n} \alpha_{ij} = 0$ for every $1 \leq i \leq m$.

Therefore, I is unique and we can take as $c(1)$ any constant $1 > c(1) > 0$. For definiteness, put $c(1) = \frac{1}{2}$.

Inductive step. Consider two cases. Denote $0 < c = \frac{1-c(d-1)}{1+(\log_2(\frac{2^d}{2^d-1})/d)} < 1 - c(d-1)$.

1. In the first case there does NOT exist a cover set $V \subset \{X_2, \ldots, X_n\}$ of variables of size $|V| \leq cn$; namely, such a set that each monomial $X_1^{\beta_1}, \ldots, X_n^{\beta_n}$

occurring in at least one of the polynomials $h_1, \ldots, h_m$ contains a variable either from V or X_1.

Let us construct sequentially a set $\{b_1, b_2, \ldots, b_l\}$ of monomials occurring in at least one of $h_1, \ldots, h_m$ such that they are pairwise disjoint in the variables and contain only the variables from $\{X_2, \ldots, X_n\}$ (so, do not contain X_1) while it is possible. Suppose that it is impossible to continue with $b_1, \ldots, b_l$ satisfying the latter conditions. Then $b_1, \ldots, b_l$ contain at most dl variables among $\{X_2, \ldots, X_n\}$, they constitute the cover set. Hence $dl \geq cn$.

Observe that for any set $\{1\} \subset I \subset \{1, \ldots, n\}$ such that $h_1^{(I)}, \ldots, h_m^{(I)}$ vanish identically the set $\{X_i, i \in I\}$ should have a common variable with each monomial $b_1, \ldots, b_l$. Therefore, the number of all such sets I does not exceed

$$2^n \left(\frac{2^d - 1}{2^d}\right)^l \leq 2^n \left(\frac{2^d - 1}{2^d}\right)^{\frac{cn}{d}} \tag{1}$$

2. In the second case such cover set V does exist. Consider any minimal $\{1\} \subset I$ such that $h_1^{(I)}, \ldots, h_m^{(I)}$ vanish identically. Denote by I_0, I_1 such sets that $\{X_i, i \in I_0\} = V \setminus \{X_i, i \in I\}, \{X_i, i \in I_1\} = V \cap \{X_i, i \in I\}$. We uniquely expand $h_j^{(I_1 \cup \{1\})} = X_1 h_{j,X_1} + \sum_{\gamma=(\cdots\gamma_i\cdots)} (\prod_{i \in I_0} X_i^{\gamma_i}) P_{j,\gamma}, 1 \leq i \leq m$, where the polynomials $P_{j,\gamma}, h_{j,X_1}$ are in the variables $X_i \notin V$. Note that $P_{j,\gamma}, h_{j,X_1}$ depend on I_0. Since V is a cover set, $\deg P_{j,\gamma} \leq d-1$ for each multiindex γ, obviously $\deg(h_{j,X_1}) \leq d-1$.

   Since $h_j^{(I)}$ vanishes identically, the polynomials $P_{j,\gamma}^{(I \setminus I_1)}$ also vanish identically, and furthermore, the polynomial $h_{j,X_1}^{(I \setminus I_1)}$ vanishes identically as well. Thus, $I \setminus I_1$ is a minimal set for which the polynomials $P_{j,\gamma}^{(I \setminus I_1)}, h_{j,X_1}^{(I \setminus I_1)}$, for all $1 \leq j \leq m$, vanish identically.

   By inductive hypothesis there are at most $2^{c(d-1)n}$ such sets $I \setminus I_1$.

   Since there are at most 2^{cn} possibilities for the sets I_0, in total we have at most

$$2^{c(d-1)n}2^{cn} \tag{2}$$

minimal sets I.

Finally, $0 < c(d) = c(d-1) + c < 1$ satisfies the inductive hypothesis for d (see (1), (2)). Thus the Proposition 3 is proved, and our Theorem follows. $\square$

References

[B83] M. Ben-Or, Lower Bounds for Algebraic Computation Trees, Proc. 15th ACM STOC (1983), pp. 80–86.

[BKL93] P. Buergisser, M. Karpinski, T. Lickteig, On Randomized Algebraic Test Complexity, J. of Complexity **9** (1993), pp. 231-251.

[F93] R. Fleischer, Decision Trees: Old and New Results, Proc. 25th ACM STOC (1993), pp. 468–477.

[FG79] F. Fussenegger, H.N. Gabow, A Counting Approach to Lower Bounds for Selection Problems, J. Assoc. Comput. Mach. 25 (1979), pp.227-238.

[GK93] D. Grigoriev, M. Karpinski, Lower Bounds on Complexity of Testing Membership to a Polygon for Algebraic and Randomized Computation Trees, Technical Report TR-93-042, International Computer Science Institute, Berkeley, 1993.

[GKV95] D. Grigoriev, M. Karpinski, N. Vorobjov, Improved Lower Bound on Testing Membership to a Polyhedron by Algebraic Decision Trees, to appear in Proc. 36th IEEE FOCS (1995).

[K73] D. E. Knuth, The Art of Computer Programming, Vol. 3, Addison-Wesley, 1973.

[MPR94] L. J. Montana, L. M. Pardo, T. Recio, A Note on Rabin's Width of a Complete Proof, Computational Complexity, 4(1994), pp. 12-36.

[R72] M. O. Rabin, Proving Simultaneous Positivity of Linear Forms, J. Comput. Sci. **6** (1972), pp. 639–650.

[SY82] J.M. Steele, A.C. Yao, Lower Bounds for Algebraic Decision Trees, J. of Algorithms, 3 (1982), pp. 1-8.

[TY93] H. F. Ting and A. C. Yao, A Randomized Algorithm for Finding Maximum with $O((\log n)^2)$ Polynomial Tests, Preprint, Princeton University, 1993.

[Y75] A. C. Yao, On the Complexity of Comparison Problems Using Linear Functions, Proc. 16th IEEE FOCS (1975), pp. 85–89.

[Y89] A. C. Yao, On Selecting the k Largest with Median Tests, Algorithmica **4** (1989), pp. 293–300.