The axiomatic derivation of absolute lower bounds

Yiannis N. Moschovakis UCLA and University of Athens

Iowa State University, April 15, 2010

The Euclidean algorithm

For
$$x,y\in\mathbb{N}=\{0,1,\ldots\},\;x\geq y\geq 1$$
,

(
$$\varepsilon$$
) $|gcd(x, y) = if (rem(x, y) = 0)$ then y else $gcd(y, rem(x, y))$

where rem(x, y) is the remainder of the division of x by y,

$$x = iq(x, y) \cdot y + rem(x, y) \quad (0 \le rem(x, y) < y)$$

 $\operatorname{calls}_{\varepsilon}(x, y) = \operatorname{the number of divisions required to compute <math>\operatorname{gcd}(x, y)$ by the Euclidean algorithm $\leq 2 \log(y) \quad (x \geq y \geq 2)$

Is the Euclidean optimal for computing gcd(x, y) from rem?

Is the Euclidean optimal for deciding coprimeness from rem?

$$x \bot y \iff \gcd(x, y) = 1$$

A partial result

Theorem (van den Dries, ynm, 2004)

If an algorithm α decides the coprimeness relation $x \perp y$ on \mathbb{N} from the primitives $\leq , +, -, iq$, rem, then for infinitely many a, b with a > b,

$$\mathsf{calls}_lpha(a,b) \ge \mathsf{depth}_lpha(a,b) > rac{1}{10} \log \log a$$
 (*)

where depth_{α}(*a*, *b*) is the least number of applications of the primitives which must be executed in sequence in the computation

- depth_{α}(x, y) is a natural parallel time complexity
- ► The result is one log short of establishing the optimality of the Euclidean (and one log = ∞ in this context)
- (*) holds for all sufficiently large a, b such that
 - $a^2 = 1 + 2b^2$ (solutions of Pell's equation), - or $a = F_{n+1}, b = F_n$ (successive Fibonacci numbers)

Claim: This applies to all algorithms from the specified primitives

Outline

Slogan: Absolute lower bound results are the undecidability facts about decidable problems

... and so their precise formulation should be a matter of logic

- (1) Tweak logic (a bit) so it applies smoothly to computation theory
- (2) Three (simple) axioms for elementary algorithms, in the style of *abstract model theory*
- (3) Verify that the axioms are satisfied by all computation models
- (4) Derive lower bounds from the axioms

Is the Euclidean algorithm optimal among its peers? (with vDD, 2004) *Arithmetic complexity* (with vDD, 2009)

Y. Mansour, B. Schieber, and P. Tiwari (1991) A lower bound for integer greatest common divisor computations, Lower bounds for computations with the floor operation

J. Meidânis (1991): Lower bounds for arithmetic problems

(Partial) algebras

Ν

• A (partial, pointed) algebra is a structure $\mathbf{M} = (M, 0, 1, \Phi^{\mathbf{M}})$

where $0, 1 \in M$, Φ is a set of function symbols (the vocabulary)

and $\Phi^{\mathsf{M}} = \{\phi^{\mathsf{M}}\}_{\phi \in \Phi}$, where each primitive $\phi^{\mathsf{M}} : M^{n_{\phi}} \rightarrow M$ is a partial function of arity n_{ϕ} determined by the symbol ϕ

$$\begin{split} \mathbf{N}_{\varepsilon} &= (\mathbb{N}, 0, 1, \text{rem}), \text{ the Euclidean algebra} \\ \mathbf{N}_{u} &= (\mathbb{N}, 0, 1, S, \text{Pd}), \text{ the unary numbers} \\ \mathbf{N}_{b} &= (\mathbb{N}, 0, 1, \text{Parity}, \text{iq}_{2}, (x \mapsto 2x), (x \mapsto 2x + 1)), \text{ the binary numbers} \\ \mathbf{N} &= (\mathbb{N}, 0, 1, +, -, \text{iq}, \text{rem}, \cdot), \text{ the full algebra of arithmetic} \\ \mathbf{I}_{\varepsilon} \upharpoonright U &= (U, 0, 1, \text{rem} \upharpoonright U) \text{ where } \{0, 1\} \subseteq U \subseteq \mathbb{N} \text{ and} \\ &(\text{rem} \upharpoonright U)(x, y) = w \iff x, y, w \in U \& \text{rem}(x, y) = w \end{split}$$

The diagram of a Φ-algebra is the set of its basic identities,

$$\mathsf{diag}(\mathsf{M}) = \{(\phi, \vec{x}, w) : \phi^{\mathsf{M}}(\vec{x}) = w\}$$

• **M** is completely determined by M, 0, 1 and diag(**M**)

Homomorphisms, embeddings and subalgebras

A homomorphism π : U → M of one Φ-algebra into another is any function π : U → M such that

$$\pi(0^{\mathsf{U}}) = y \iff y = 0^{\mathsf{M}}, \quad \pi(1^{\mathsf{U}}) = y \iff y = 1^{\mathsf{M}},$$

and for all $\phi \in \Phi, x_1, \ldots, x_n, w \in U$,

$$\phi^{\mathsf{U}}(x_1,\ldots,x_n)=w\implies\phi^{\mathsf{M}}(\pi x_1,\ldots,\pi x_n)=\pi w$$

- π is an embedding if, in addition, it is one-to-one
- Subalgebras:

We will use finite subalgebras U ⊆_p M to represent calls to the primitives executed during a computation in M Terms and depth

The terms of a vocabulary Φ : $t := 0 | 1 | v | \phi(t_1, \ldots, t_n)$



$${\sf depth}(0) = {\sf depth}(1) = {\sf depth}(v) = 0,$$

 ${\sf depth}(\phi(t_1, \ldots, t_n)) = {\sf max}({\sf depth}(t_1), \ldots, {\sf depth}(t_n)) + 1$

Term evaluation and algebra generation

• If $\vec{v} = v_1, \ldots, v_n$ includes all the variables which occur in t,

 $t(\vec{v}) = (t,\vec{v})$

▶ For any subalgebra $\mathbf{U} \subseteq_{p} \mathbf{M}$ and any $\vec{x} = x_1, \ldots, x_n \in U$,

 $t^{U}[\vec{x}] =$ the value of t in **U** when $\vec{v} := \vec{x}$ (if it converges)

•
$$G_m[\mathbf{U}, \vec{x}] = \{ t^{\mathbf{U}}[\vec{x}] : \operatorname{depth}(t(\vec{v})) \leq m \& t^{\mathbf{U}}[\vec{x}] \downarrow \}$$

- $G_{\infty}[\mathbf{U}, \vec{x}] = \bigcup_{m} G_{m}[\mathbf{U}, \vec{x}] = \text{the subalgebra of } \mathbf{U} \text{ generated by } \vec{x}$
- depth $(w, \mathbf{U}, \vec{x}) = \min\{m : w \in G_m[\mathbf{U}, \vec{x}]\}$

depth (w, \mathbf{U}, \vec{x}) is the least number of applications of the primitives which must be executed in sequence to construct w from \vec{x} in **U**

$$\bullet \quad \mathsf{depth}_{\vec{x}}(\mathbf{U}) = \mathsf{max}\{\mathsf{depth}(w,\mathbf{U},\vec{x}) : w \in U\} \quad (\mathbf{U} \text{ generated by } \vec{x})$$

The depth complexity of values of a function

Basic principle: If an algorithm α computes $f : M^n \to M$ from the primitives of \mathbf{M} , then $f(\vec{x}) \in G_{\infty}[\mathbf{M}, \vec{x}]$ and

 $\mathsf{calls}_\alpha(\vec{x}) \geq \mathsf{depth}(f(\vec{x}), \mathbf{M}, \vec{x})$

- The value must be constructed by the primitives from the input
- Can be used to derive lower bounds for functions which grow fast, e.g., multiplication or the Ackermann function (in unary or binary arithmetic)

We can also exploit gaps in $G_m[\mathbf{M}, a]$ when a is large compared to m:

$$G_m[\mathbf{N}_b,a]: 0 \ 1 \ 2 \ \cdots \ 2^{m+1}-1 \ \ \mathsf{gap} \ \ \mathsf{iq}(a,2^m)\cdots \ a \ \cdots$$

Theorem (van den Dries)

If an algorithm α computes gcd(x, y) from

 $\begin{array}{rll} + & \dot{-} & < & = & {\rm iq \ rem } & \cdot \\ then \ for \ all \ a > b \ such \ that \ a^2 = 1 + 2b^2 \ ({\rm Pell \ pairs}) \\ {\rm calls}_{\alpha}(a+1,b) \geq {\rm depth}(a+1,b) \geq \frac{1}{4}\sqrt{\log\log b} \end{array}$

This is the best known lower bound for the gcd from primitives which include multiplication

This method cannot yield lower bounds for decision problems
 (because their value (0 or 1) is available with no computation)

I The Locality Axiom

An algorithm α of arity n of an algebra $\mathbf{M} = (M, 0, 1, \Phi^{\mathbf{M}})$ assigns to each subalgebra $\mathbf{U} \subseteq_{p} \mathbf{M}$ an n-ary (strict) partial function

$$\overline{\alpha}^{\mathsf{U}}: U^n \rightharpoonup U$$

An M-algorithm α "computes" a partial function α^M : Mⁿ → M using the primitives of M as oracles, and it can be naturally localized (restricted) to arbitrary subalgebras of M

We write

$$\mathbf{U}\vdash\overline{\alpha}(\vec{x})=w\iff\vec{x}\in U^n,w\in U\text{ and }\overline{\alpha}^{\mathbf{U}}(\vec{x})=w$$

II The Homomorphism Axiom

If α is an n-ary algorithm of \mathbf{M} , $\mathbf{U}, \mathbf{V} \subseteq_{p} \mathbf{M}$, and $\pi : \mathbf{U} \rightarrow \mathbf{V}$ is a homomorphism, then

$$\mathbf{U}\vdash\overline{\alpha}(\vec{x})=w\implies\mathbf{V}\vdash\overline{\alpha}(\pi\vec{x})=\pi w\quad(x_1,\ldots,x_n,w\in U)$$

In particular, if $\mathbf{U} \subseteq_{p} \mathbf{M}$, then $\overline{\alpha}^{\mathbf{U}} \sqsubseteq \overline{\alpha}^{\mathbf{M}}$

When asked for φ^U(x), the oracle for φ may consistently provide φ^V(πx), if π is a homomorphism.

III The Finiteness Axiom

If α is an n-ary algorithm of **M**, then

 $\mathbf{M}\vdash\overline{\alpha}(\vec{x})=w$

 \implies there is a finite $\mathbf{U} \subseteq_p \mathbf{M}$ generated by \vec{x} such that $\mathbf{U} \vdash \overline{\alpha}(\vec{x}) = w$

- ► The algorithm must construct the arguments *u* for every call φ(*u*) to the primitives, and so the entire computation takes place within the subalgebra generated by the input *x*
- ► A computation of a^M(x) from the primitives of M is finite, and so it makes finitely many calls to the primitives
- Intuitively, the axiom is satisfied by any U whose diagram includes (φ, ū, φ^M(ū)) for every call to φ^M made by α during the computation

All elementary algorithms satisfy I - III (with suitable **M**)

- Explicit computation: $\overline{\alpha}^{M}(\vec{x}) = t^{M}[\vec{x}]$, where $t(\vec{v})$ is a Φ -term
- *α*^M is the partial function computed a fixed recursive (McCarthy) program *P* in the vocabulary Φ
- ▶ $\overline{\alpha}^{M}$ is computed in PCF (typed λ -calculus) above the algebra **M**
- $\blacktriangleright \ \overline{\alpha}^{\mathbf{M}}$ by computed by non-deterministic versions of any of these

Note: For computation models (e.g., Turing machines), the functions built into the model must be included among the primitives of M.

Abstract algorithms

I, Locality Axiom: An abstract algorithm α of arity n of an algebra M = (M, 0, 1, Φ^M) assigns to each subalgebra U ⊆_p M an n-ary partial function α^U : Uⁿ → U

$$\mathbf{U}\vdash\overline{\alpha}(\vec{x})=w\iff\overline{\alpha}^{\mathbf{U}}(\vec{x})=w$$

II, Homomorphism Axiom: If U, V ⊆_p M, and π : U → V is a homomorphism, then

$$\mathbf{U}\vdash\overline{\alpha}(\vec{x})=w\implies\mathbf{V}\vdash\overline{\alpha}(\pi\vec{x})=\pi w$$

Set $\mathbf{U} \vdash_{g} \overline{\alpha}(\vec{x}) = w \iff \mathbf{U}$ is generated by $\vec{x} \& \mathbf{U} \vdash \overline{\alpha}(\vec{x}) = w$

▶ III, Finiteness Axiom:

$$\mathbf{M} \vdash \overline{\alpha}(\vec{x}) = w \implies$$
 there is a finite $\mathbf{U} \subseteq_p \mathbf{M}$ such that $\mathbf{U} \vdash_g \overline{\alpha}(\vec{x}) = w$

Complexity functions for abstract algorithms

Suppose α is an *n*-ary abstract algorithm of **M** and $\mathbf{M} \vdash \overline{\alpha}(\vec{x}) = w$

$$| \operatorname{calls}_{\alpha}(\vec{x}) = \min\{ |\operatorname{diag}(\mathbf{U})| : \mathbf{U} \vdash_{g} \overline{\alpha}(\vec{x}) = w \}$$

(the least number of calls α must execute to compute $\overline{\alpha}^{M}(\vec{x})$)

$$\bullet |\operatorname{size}_{\alpha}(\vec{x}) = \min\{|U \setminus \{0, 1, \vec{x}\}| : \mathbf{U} \vdash_{g} \overline{\alpha}(\vec{x}) = w\}$$

(the least number of elements of **M** (other than $0, 1, \vec{x}$) that α must see to compute $\overline{\alpha}^{M}(\vec{x})$)

$$\bullet | \mathsf{depth}_{\alpha}(\vec{x}) = \mathsf{min}\{\mathsf{depth}_{\vec{x}}(\mathsf{U}) : \mathsf{U} \vdash_{g} \overline{\alpha}(\vec{x}) = w\}$$

(the least number of calls α must execute in sequence to compute $\overline{\alpha}^{M}(\vec{x})$)

$$\mathsf{Thm} \, \left| \, \mathsf{depth}_{\alpha}(\vec{x}) \leq \mathsf{size}_{\alpha}(\vec{x}) \leq \mathsf{calls}_{\alpha}(\vec{x}) \right| \\$$

These notions agree with standard definitions for concrete algorithms

★ The certification relation $\mathbf{U} \Vdash f(\vec{x}) = w$

Suppose $f: M^n \to M$, $\mathbf{U} \subseteq_{p} \mathbf{M}$.

• A homomorphism $\pi: \mathbf{U} \to \mathbf{M}$ respects f at \vec{x} if

$$\vec{x} \in U \& f(\vec{x}) \in U \& \pi(f(\vec{x})) = f(\pi(\vec{x}))$$

 $|\mathbf{U} \Vdash f(\vec{x}) = w \iff$ every homomorphism $\pi : \mathbf{U} \to \mathbf{M}$ respects f at \vec{x}

Lemma

(1) If an abstract **M**-algorithm α computes $f: M^n \to M$, then

$$\mathbf{U}\vdash_{g}\overline{\alpha}(\vec{x})=w\implies \mathbf{U}\Vdash f(\vec{x})=w$$

(2) If some abstract **M**-algorithm α computes $f : M^n \to M$ and $f(\vec{x}) = w$, then there is a finite **U** such that **U** $\Vdash f(\vec{x}) = w$ Proof. By the Homomorphism and Finiteness axioms.

Certificates

For coprimeness in the Euclidean algebra $N_{\varepsilon} = (\mathbb{N}, 0, 1, \text{rem})$,

$$\mathbf{U} = \{ \operatorname{rem}(x, y) = r_1, \operatorname{rem}(y, r_1) = r_2, \dots, \operatorname{rem}(r_n, r_{n+1}) = 1 \}$$

certifies that $x \perp y$.

To prove (by this method) that the Euclidean is number-of-calls optimal for coprimeness, we would need to show that every certificate for $x \perp y$ in \mathbf{N}_{ε} is at least as large as $n + 2 = \text{calls}_{\varepsilon}(x, y)$ for infinitely many (x, y).

Complexity functions for computable functions

Suppose some abstract **M**-algorithm α computes $f: M^n \rightarrow M$.

► calls_f(
$$\mathbf{M}, \vec{x}$$
) = min{|diag(\mathbf{U})| : $\mathbf{U} \Vdash f(\vec{x}) = w$ } ≤ calls _{α} (\vec{x})

► size_f(
$$\mathbf{M}, \vec{x}$$
) = min{ $|U \setminus \{0, 1, \vec{x}\}| : \mathbf{U} \Vdash f(\vec{x}) = w$ } ≤ size _{α} (\vec{x})

▶ depth_f(\mathbf{M}, \vec{x}) = min{depth_{\vec{x}}(\mathbf{U}) : $\mathbf{U} \Vdash f(\vec{x}) = w$ } ≤ depth_{α}(\vec{x})

Thm depth_f(
$$\mathbf{M}, \vec{x}$$
) \leq size_f(\mathbf{M}, \vec{x}) \leq calls_f(\mathbf{M}, \vec{x})

Interpretation: depth_f(\mathbf{M}, \vec{x}) is the least number of calls to the primitives which must be executed in sequence by any algorithm which computes $f : M^n \to M$ from the primitives of \mathbf{M}

► These complexities are most useful for relations R ⊆ Mⁿ whose characteristic functions take values in {0,1}

Outline of a proof

Theorem (van den Dries, ynm)

$$In \mathbf{M} = (\mathbb{N}, 0, 1, \leq, +, -, \text{iq}, \text{rem}): \text{ if a is sufficiently large and}$$

$$\boxed{a^2 = 1 + 2b^2} \text{ or } \boxed{a = F_{n+1}, b = F_n}, \text{ then}$$

$$\operatorname{depth}_{\underline{\mathbb{H}}}(\mathbf{M}, a, b) > \frac{1}{10} \log \log(a) \qquad (*)$$

So if α decides coprimeness in **M**, then (*) holds with depth_{α}(a, b)

Pf. If $2^{2^{4m+6}} \leq a$, then every $X \in G_m[\mathbf{M}, a, b]$ can be written uniquely as $X = \frac{x_0 + x_1 a + x_2 b}{x_3} \quad \text{with } x_i \in \mathbb{Z}, \quad |x_i| \leq 2^{2^{4m}}$

and we can define $\pi: \mathit{G}_m[\mathbf{M}, a, b] \rightarrowtail \mathbf{M}$ letting $\lambda = 1 + a!$,

$$\pi(X) = \frac{x_0 + x_1\lambda a + x_2\lambda b}{x_3}, \text{ so } (\pi(a), \pi(b)) = (\lambda a, \lambda b)$$

The "universal constant" $\frac{1}{10}$

Detailed version of result

In $\mathbf{M} = (\mathbb{N}, 0, 1, \leq, +, -, iq, rem)$: if $1 < \xi < 2$, ξ is a quadratic algebraic irrational, C > 0, a is sufficiently large, and a $\bot b$, then

$$rac{1}{Cb^2} < \left| \xi - rac{a}{b}
ight| < rac{1}{b^2} \implies ext{depth}_{\perp}(a,b) > rac{1}{K} \log \log a,$$

with $K \geq 2\log(\log C + 19)$

- Liouville: for sufficiently large C, infinitely many a, b satisfy the hypothesis
- With $\xi = \sqrt{2}$ and $a^2 = 1 + 2b^2$, we can take $\mathcal{C} = 5, \mathcal{K} \geq 10$
- ▶ With $\xi = \frac{1}{2}(1 + \sqrt{5})$ and $a = F_{n+1}, b = F_n$, we can take again, $C = 5, K \ge 10$

 $\textbf{M} = (\mathbb{N}, 0, 1, \mathsf{Parity}, \mathsf{iq}_2, \leq, +, -, \mathsf{Presburger\ functions})$

• (van den Dries, ynm) If R(x) is one of the relations

then for some r > 0 and infinitely many a,

 $\operatorname{depth}_R(\mathbf{M}, a) > r \log(a)$

▶ (van den Dries, ynm) For some r > 0 and infinitely many a, b,

 $depth_{\parallel}(\mathbf{M}, a, b) > r \log(max(a, b))$

▶ (Joe Busch) If $R(x, p) \iff x$ is a square mod p, then for some r > 0 and a sequence (a_n, p_n) with $p_n \to \infty$,

$$depth_R(\mathbf{M}, a_n, p_n) > r \log(p_n)$$

In the last two examples, the results match up to a multiplicative constant well-known known binary algorithms, so these are optimal

Non-uniform complexity

What if you are only interested in deciding $R(\vec{x})$ for *n*-bit numbers $(< 2^n)$ and you are willing to use a different algorithm for each *n*?

- ▶ The lookup algorithm: For any *k*-ary relation *R* on \mathbb{N} and each *n*, there is an **N**_b-term (with conditionals) $t_n(\vec{v})$ of depth $\leq n = \log(2^n)$ which decides $R(\vec{x})$ for all $\vec{x} < 2^n$.
- Non-uniform lower bounds on depth are never greater than log
- The best ones establish the optimality of the lookup algorithm and are most interesting when some uniform algorithm matches the lookup algorithm up to a multiplicative constant
- They are quite easy for Presburger primitives

Coprimeness from division, non-uniformly

Theorem (van den Dries, ynm) Let $\mathbf{M} = (\mathbb{N}, 0, 1, \leq, +, -, iq, rem)$ and for each n, let

$$x_{\perp}y \iff x, y < 2^n \& \operatorname{gcd}(x, y) = 1.$$

There is some r > 0, such that for all sufficiently large n, there are a, $b < 2^n$ such that

$$\mathsf{calls}_{\underline{\parallel}_n}(\mathbf{M}, a, b) \ge \mathsf{size}_{\underline{\parallel}_n}(\mathbf{M}, a, b) > r \log n$$
 (**)

So if n is large enough and α decides coprimeness in **M** for all $x, y < 2^n$, then (**) holds with calls_{α}(a, b), size_{α}(a, b) on the left

I do not know how to get the corresponding result for depth_⊥

Concluding remarks

- (1) A technique for deriving lower bounds for decision problems which are absolute, i.e., they hold of all computational models
- (2) Main limitation: in its current version, it only yields lower bounds which are no better than O(n) (linear in the length of the input)
- (3) Problem: prove that the Euclidean algorithm is optimal for computing the gcd in the algebra $\mathbf{N}_{\varepsilon} = (\mathbb{N}, 0, 1, \text{rem})$
- (4) *Problem*: prove an $O(n^2)$ lower bound for *primality* in $\mathbf{N}_b = (\mathbb{N}, 0, 1, \text{Parity}, \text{iq}_2, (x \mapsto 2x), (x \mapsto 2x + 1))$

Comment: (4) may need some number theory, but it will also need some logical analysis of computation (since the entire input is known in O(n) steps)