Small circuits and dual weak PHP in the universal theory of p-time algorithms

Jan Krajíček

Faculty of Mathematics and Physics
Charles University*

Abstract

We prove, under a computational complexity hypothesis, that it is consistent with the true universal theory of p-time algorithms that a specific p-time function extending n bits to $m \geq n^2$ bits violates the dual weak pigeonhole principle: every string $y \in \{0,1\}^m$ equals the value of the function for some $x \in \{0,1\}^n$. The function is the truth-table function assigning to a circuit the table of the function it computes and the hypothesis is that every language in P has circuits of a fixed polynomial size n^d.

Consider a first-order language having a function symbol for every deterministic p-time clocked Turing machine, the symbol being interpreted over $\mathbf{N}$ by the function computed by the machine. Relations computable by p-time machines are formally represented by their characteristic functions but we write, for example, $x \leq y$ instead of $\leq (x,y) = 1$. We shall denote this language $L_{\rm PV}$ and the theory of all true universal sentences in the language by $T_{\rm PV}$; the notation alludes to the influential theory PV introduced by Cook [2] although its language is defined in a much more complicated way (because of the intended links with proof complexity, cf.[2] or [8, Chpt.5]). Note that a number of important theorems from computational complexity, including the PCP theorem or all valid instances of the NP-completeness of SAT (where there are p-time functions sending witnesses to witnesses) or various lower bounds, can be expressed as universal statements in the language and hence are axioms of theory $T_{\rm PV}$ (cf. [13, Sec.22.3]).

Dual weak pigeonhole principle (dWPHP) for a function g says that for no $n < m$ can g map n-bit strings onto all m-bit strings. It was first considered in the context of bounded arithmetic by Wilkie who proved a witnessing theorem for a particular theory having dWPHP among its axioms; this was written up in [8, Thm.7.3.7]. The theory was suggested as the *basic theory* (BT) for formalization of complexity theory (and, in particular, of probabilistic constructions) in [9] and indeed Jeřábek [3, 4, 5] succeeded in it spectacularly. The dWPHP is

*Sokolovská 83, Prague, 186 75, The Czech Republic, krajicek@karlin.mff.cuni.cz

also linked with proof complexity (cf. [10]) and one can view the question (posed in [9]) whether one of the theories related to p-time algorithms proves dWPHP for p-time functions as a uniform version of particular propositional lengths-of-proofs problems (about proof complexity generators, cf.[13, Sec.19.4]). We do not give definitions of the notions mentioned above or details of the statements as they are not technically relevant to this paper and serve here only as a motivation for the expert reader. The non-expert reader can find this background in the references given and, in particular, all of it in [13].

The unprovability of dWPHP in T^1_2 implies that P $\neq$ NP: Paris, Wilkie and Woods [15] proved that dWPHP for p-time functions is provable in a theory having induction for all predicates in the p-time hierarchy (theory T_2(PV) of Buss [1]) but that theory would follow from T^1_2 if it were that P = NP. This is because if satisfiability can be solved by p-time algorithm f that statement is universal:

$$Sat(x, y) \rightarrow Sat(x, f(x))$$

and hence in T^1_2 (here $Sat(x, y)$ formalizes that y is a satisfying assignment for formula x). But then every bounded formula is provably in T^1_2 equivalent to an open formula and hence induction for all bounded formulas follows from induction for open formulas which is in T^1_2 provable via usual binary search. This means that if we want to prove that dWPHP is not provable in T^1_2 we ought to expect to use some hypothesis that itself implies P $\neq$ NP. The hypothesis we shall use is the following:

Hypothesis (H):

There exists constant $d \geq 1$ such that every language in P can be decided by circuits of size $O(n^d)$: P $\subseteq$ Size(n^d).

The popular expert opinion finds (H) unlikely, I suppose, but the reader should note that there are no technical results that would support the skepticism. In anything, the notorious unability to prove even $10n$ lower bound for general circuits suggest that the possibility that (H) is true cannot be simply dismissed. It has also the attractive feature that it implies P $\neq$ NP (there are languages in the polynomial-time hierarchy that have no size $O(n^d)$ circuits, cf. Kannan's theorem [7]). Thus, in principle, one could prove P $\neq$ NP by proving circuit upper bounds rather than by proving lower bounds. It is less attractive that it implies also E $\subseteq$ Size($2^{o(n)}$) (a language in E becomes p-time computable if the inputs are padded) and hence it disproves the foundational hypothesis of universal derandomization. But this is not an a priori reason to abandon (H) as E $\not\subseteq$ Size($2^{o(n)}$) is itself only a hypothesis. On the other hand (H) is good for proof complexity: together with [11, Thm.2.1] the statement E $\subseteq$ Size($2^{o(n)}$) (and hence (H)) imply that either NP $\neq$ coNP or that there is no p-optimal propositional proof system; proving (or disproving) one of these two statements are the two fundamental problems of proof complexity. The hypothesis (with linear size circuits) is often attributed to Kolmogorov, see the discussion in [6, Sec.20.2].

Next we need to define the **truth-table function $\mathbf{tt}_{s,k}$**. It takes as an input

a circuit with k inputs of size $\leq s$ and outputs its truth table, 2^k bits. A size $\leq s$ circuit can be encoded by, say, $10s \log s$ bits exactly and hence for $10s \log s < 2^k$ this is a function from a smaller set into a bigger one. Our size function $s = s(k)$ will have the form $s(k) := 2^{\epsilon k}$ for some fixed $0 < \epsilon < 1$. Hence any such $\mathbf{tt}_{s,k}$ is a p-time function.

Theorem 1

Assume hypothesis (H). Then for every $0 < \epsilon < 1$ and $s = s(k) := 2^{\epsilon k}$ the theory T_{PV} does not prove the sentence

$$\forall 1^m (m = 2^k > 1) \exists y \in \{0,1\}^m \forall x \in \{0,1\}^n,\ \mathbf{tt}_{s,k}(x) \neq y \tag{1}$$

expressing the dWPHP for $\mathbf{tt}_{s,k}$, where $n := 10s \log s$.

Proof :

Assume that T_{PV} proves (1). By the KPT theorem (cf.[14] or [8, Thm.7.4.1] or [13, Cor.12.2.4]) there a p-time functions

$$f_1(z), f_2(z, w_1), \ldots, f_t(z, w_1, \ldots, w_{t-1}) \tag{2}$$

such that for any $m = 2^k > 1$ and any $b_1, \ldots, b_t, C_1, \ldots, C_{t-1}$:

- either $b_1 \notin rng(\mathbf{tt}_{s,k})$ for $b_1 = f_1(1^m) \in \{0,1\}^m$ or, if $b_1 \in rng(\mathbf{tt}_{s,k})$ and $b_1 = \mathbf{tt}_{s,k}(C_1)$,
- $b_2 \notin rng(\mathbf{tt}_{s,k})$ for $b_2 = f_2(1^m, C_1) \in \{0,1\}^m$ or, if $b_2 \in rng(\mathbf{tt}_{s,k})$ and $b_2 = \mathbf{tt}_{s,k}(C_2)$,
- ..., or
- $b_t \notin rng(\mathbf{tt}_{s,k})$ for $b_t = f_t(1^m, C_1, \ldots, C_{t-1}) \in \{0,1\}^m$.

Define constants $\delta_i := (2d)^{-i}$, for $i = 0, \ldots, t$, and parameters $m_i := m^{\epsilon \delta_i}$ where d is the constant from (H) and m is large enough.

We first show that f_1 cannot find a suitable b_1. Define the function $\hat{f}_1$ that has $m_t + k$ variables and on inputs 1^{m_t} and $i \in \{0,1\}^k$ computes the i-th bit of $f_1(1^m)$. The string 1^{m_t} has the only purpose to make $\hat{f}_1$ p-time. By hypothesis (H) there is a circuit $C_1'(z, i)$ with the same variables as $\hat{f}_1$ that computes $\hat{f}_1$. Define C_1 by substituting 1^{m_t} for z in C_1' and leaving just the k variables for bits of i. Note that C_1 has size $O((m_t + k)^d)$ and thus can be encoded by $\leq m_{t-1}$ bits. Further, by its definition, $\mathbf{tt}_{s,k}(C_1) = b_1$.

Now we show that f_2 does not compute a suitable $b_2 := f_2(1^m, C_1)$ either. As before define function $\hat{f}_2$ that now takes three inputs: string $1^{m_{t-1}}$, circuit C_1 (substituted for w_1) and $i \in \{0,1\}^k$, and computes the i-th bit of $f_2(1^m, C_1)$. Applying (H) we get a circuit C_2' with the same $2m_{t-1} + k$ variables as $\hat{f}_2$ that computes the function. Define C_2 by substituting $1^{m_{t-1}}$ for z and bits defining C_1 for w_1 in C_2', and leaving just the k variables for bits of i. Note that C_2 can be encoded by $\leq m_{t-2}$ bits and $\mathbf{tt}_{s,k}(C_2) = b_2$.

Continuing in an analogous way for t steps we show that the t-tuple of functions (2) cannot have the claimed property. Note that the final C_t witnessing that $b_t \in rng(\mathbf{tt}_{s,k})$ too can be encoded by m_0 bits and hence all circuits C_i have size at most $m_0 = m^\epsilon = 2^{\epsilon k}$.

q.e.d.

The reader who is confident that hypothesis (H) is false can interpret the theorem as saying that in order to disprove (H) it suffices to prove in $T_{\rm PV}$ the existence of (a table of) a Boolean function with an exponential circuit complexity.

It would be desirable to prove the theorem for some other p-time function g under a weaker or different hypothesis than (H). A good candidate for g may be the proof complexity generator defined in [12, Sec.3] (or see [13, Sec.19.4]). It is not difficult to modify the proof of Theorem 1 for this function (with suitable parameters).

References

[1] S. R. Buss, *Bounded Arithmetic*. Naples, Bibliopolis, (1986).

[2] S. A. Cook, Feasibly constructive proofs and the propositional calculus, in: *Proc. 7^{th} Annual ACM Symp. on Theory of Computing* (STOC), (1975), pp. 83-97. ACM Press.

[3] E. Jeřábek, *Weak pigeonhole principle, and randomized computation*, Ph.D. thesis, Charles University, Prague, (2005).

[4] E. Jeřábek, Dual weak pigeonhole principle, Boolean complexity, and derandomization, *Annals of Pure and Applied Logic*, **129**, (2004), pp.1-37.

[5] E. Jeřábek, Approximate counting in bounded arithmetic, *J. of Symbolic Logic*, **72(3)**, (2007), pp.959-993.

[6] S. Jukna, *Boolean function complexity*, Springer, 2012.

[7] R. Kannan, Circuit-size lower bounds and non-reducibility to sparse sets, *Information and Control*, **55(13)**, (1982), pp.4056.

[8] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Encyclopedia of Mathematics and Its Applications, Vol. **60**, Cambridge University Press, (1995).

[9] J. Krajíček, On the weak pigeonhole principle, *Fundamenta Mathematicae*, Vol.**170(1-3)**, (2001), pp.123-140.

[10] J. Krajíček, Dual weak pigeonhole principle, pseudo-surjective functions, and provability of circuit lower bounds, *J. of Symbolic Logic*, **69(1)**, (2004), pp.265-286.

[11] J. Krajíček, Diagonalization in proof complexity, *Fundamenta Mathematicae*, **182**, (2004), pp.181-192.

[12] J. Krajíček, A proof complexity generator, in: *Proc. from the 13th Int. Congress of Logic, Methodology and Philosophy of Science (Beijing, August 2007)*, King's College Publications, London, ser. Studies in Logic and the Foundations of Mathematics. Eds. C.Glymour, W.Wang, and D.Westerstahl, (2009), pp.185-190.

[13] J. Krajíček, *Proof complexity*, Encyclopedia of Mathematics and Its Applications, Vol. **170**, Cambridge University Press, to appear in 2019.

[14] J. Krajíček, P. Pudlák and G. Takeuti, Bounded arithmetic and the polynomial hierarchy, *Annals of Pure and Applied Logic*, **52**, (1991), pp.143–153.

[15] J. Paris, A. J. Wilkie and A. Woods, Provability of the Pigeonhole Principle and the Existence of Infinitely Many Primes, *Journal of Symbolic Logic*, **53(4)**, (1988), pp.1235-1244.