Warsaw, December 13, 2012

Dr Leszek Kołodziejczyk
Institute of Mathematics
University of Warsaw
Banacha 2
02-097 Warsaw, Poland

Report on Sebastian Müller's doctoral dissertation
"On the power of weak extensions of $\mathbf{V^0}$"

Sebastian Müller's dissertation contributes to the twin research areas of bounded arithmetic and propositional proof complexity. Propositional proof complexity, the study of lengths of proofs in various proof systems for propositional logic, is the better-known of the two, due to its important position in theoretical computer science. It is a *bona fide* subfield of computational complexity theory, since the main goal of proof complexity, obtaining superpolynomial lower bounds on proof size for all possible proof systems, amounts to showing that the complexity class NP is not closed under complement. Moreover, proof complexity is also relevant to the theory of algorithms solving the propositional satisfiability problem (SAT-solvers), because the run of such an algorithm on an unsatisfiable formula can be viewed as the proof of its negation in a certain proof system.

Weak axiomatic theories of arithmetic, of the kind later known as bounded arithmetic, were first studied in the 1970's with a motivation which was at least in part philosophical. Nowadays, bounded arithmetic is, among other things, the source of some indispensable techniques used in proof complexity. The connection between the two fields is based on a correspondence, or translation, between theories of arithmetic and propositional proof systems. The provability of an arithmetical statement in a given theory implies the existence of short proofs of related propositional tautologies in the proof system associated with the theory. In practice, this is especially helpful in proving upper bounds on the length of propositional proofs, as the proofs are often much easier to find or describe on the first-order than on the propositional level. Additionally, the arithmetic setting makes it possible to employ the tools of first-order model theory.

The work of S. Müller presented in the thesis exploits precisely this connection between arithmetic theories and propositional proof systems. The proofs of the main results share a similar outline: a certain statement or class of statements is shown to be provable in a bounded arithmetic theory; by the correspondence between theories and proof systems, this means that some tautologies have short proofs in a given system; depending on the context, this provides additional structural information on the propositional level (a separation or simulation theorem for some proof systems). Most of the technical work happens within the arithmetic theory, though typically it is the propositional result which is the main goal. The focus is on some extensions of the very weak two-sorted theory $\mathbf{V^0}$, which on the propositional level corresponds to the extremely important class of *constant-depth* proof systems (i.e. systems using only formulas in which the number of alternations between $\wedge$ and $\vee$ is bounded by a constant).

The results of the thesis. The dissertation contains five sections and three appendices. Its core part consists of Sections 2 to 4, which are devoted to the presentation of

the research results. In Sections 2 and 3, most of the proofs are only briefly sketched, with the technical details relegated to the appendices (Appendices A and B, respectively). A side effect is that the appendices make up almost two thirds of the thesis.

Section 2, and the corresponding appendix A, describe joint work of the author and Iddo Tzameret on refutations of *random 3-CNF* formulas. A random 3-CNF with n variables and clause-to-variable ratio Δ is obtained by selecting independently at random $m = \Delta n$ three-literal clauses with variables from among $x_1, \ldots, x_n$ (w.r.t. the uniform distribution, with repetitions). It is well-known that if Δ is above a certain constant, random 3-CNFs become unsatisfiable with high probability. However, the provability of this in specific proof systems is another matter. Random CNFs with Δ sufficiently close to the unsatisfiability threshold are conjectured to be hard to refute even in strong proof systems, whereas the weak but non-trivial Resolution system is unable to refute them efficiently even for $\Delta = n^{0.5-\epsilon}$.

The author and I. Tzameret prove (Theorem 2.1 in the thesis) that if Δ is at least $cn^{0.4}$, c sufficiently large, then random 3-CNFs have polynomial-size refutations in constant-depth systems in a language with $\wedge, \vee, \neg$ and additional threshold connectives, which evaluate to True if and only if the number of their True inputs exceeds a certain threshold. This is a new and interesting result, as until now polynomial upper bounds on refutation size of random 3-CNFs in systems of modest strength were only known for much higher Δ (almost on the order of n). The new upper bound also shows that the known exponential separation between proof size in Resolution and in constant depth systems with threshold connectives holds what in could be called an "average-case" setting. One could argue that this suggests the desirability of developing SAT-solvers based on systems stronger than Resolution.

The proof of the new upper bound is based on a result of Feige *et al.* that if $\Delta = cn^{0.4}$ or higher, w.h.p. there exists a small witness for the unsatisfiability of a random 3-CNF. Müller and Tzameret make the ingenious observation that to obtain the upper bound it suffices to formalize the correctness proof for these witnesses in the arithmetic theory $\mathbf{VTC^0}$, an extension of $\mathbf{V^0}$ corresponding to constant depth proofs with threshold connectives. The formalization, while not exceedingly difficult, is delicate, as the correctness proof given by Feige *et al.* makes extensive use of linear algebra over the reals. $\mathbf{VTC^0}$ cannot reason about reals, so rational approximations have to be used, and it (conjecturally) knows rather little about linear algebra, so various objects whose existence is not provable have to be dealt with by means of quantification. Müller and Tzameret handle these difficulties quite well.

Section 3, and the corresponding appendix B, turn to the study of more typical textbook proof systems, often called Frege systems, and the theory $\mathbf{VNC^1}$ corresponding to them. The author uses an argument in the spirit of the proof of Nepomnjaščiĭ's theorem to show that polylogarithmic cuts in models of $\mathbf{V^0}$ actually satisfy $\mathbf{VNC^1}$, a provably stronger theory (Theorem 3.7). By the correspondence between arithmetic and propositional logic, and a twist on a standard argument involving provability of so-called reflection principles for proof systems, this implies that constant-depth systems subexponentially simulate Frege systems, in the sense that for each $D \geq 1, \delta > 0$ there exists some d such that for all formulas of depth D, the existence of a Frege proof of size n implies the existence of a depth d proof of size $O(2^{n^\delta})$ (Theorem 3.8). As the author acknowledges, this simulation result was first proved by Filmus *et al.*, but the proof given

here is considerably shorter and (to my taste) more elegant, and the underlying method perhaps more widely applicable. The downside of the author's method is that the result is slightly weaker: the explicit syntactic transformations performed by Filmus *et al.* yield some quantitative information linking D, d and δ, which is lost in the arithmetic-based approach.

As in Section 2, also here it is the arithmetic part which makes up the bulk of the proof. It has to be pointed out that this part is not altogether new. It is implicit in a Nepomnjaščiĭ-style proof appearing in a paper by D. Zambella (*End extensions of models of linearly bounded arithmetic*, Ann. Pure App. Logic 88 (1997), 263-277) that polylog cuts of models of $\mathbf{V^0}$ satisfy the theory Σ^p_0-**Rec**, nowadays known to coincide with **VL**, an extension of $\mathbf{VNC^1}$. The author seems unaware of this. Perhaps he should not be blamed, since Zambella's theorem is stated in terms of end-extensions (of models of the theory $I\Delta_0$) rather than cuts, and the proof as it stands applies only to certain "canonical" models of $\mathbf{V^0}$, though it is readily seen to generalize to all others. Zambella's paper does not deal with propositional logic at all, but it seems clear that Müller's technique for proving subexponential simulations should also work for the slight extension of Frege systems corresponding to **VL**.

At the end of Section 3, there is a brief subsection (3.2) without a corresponding part in the technical Appendix B, purporting to prove that the theory of polylog cuts in models of $\mathbf{V^0}$ contains not just $\mathbf{VNC^1}$, but the presumably stronger theory **VNC**. Unfortunately, I am not convinced that the argument here is correct. It is based on an algorithm claimed to evaluate so-called NC^k circuits, $k \geq 1$, in polynomial time and polylogarithmic space. However, I do not see why the algorithm "clearly runs in polynomial time": it appears to need quasipolynomial time, whereas natural modifications which would run in polynomial time would also require more space. In fact, proving the existence of an algorithm with the properties claimed by the author would solve the long-standing and apparently still open problem whether the uniform version of the complexity class NC is contained in SC (simultaneous polynomial time and polylog space).

Section 4 contains some preliminary work towards a conditional proof that the Resolution system does not have a desirable feature known as automatizability. In an approach the author envisions (-ed?), one step would require a strengthening of the results of Section 3, and another step would be to prove the correctness of the Diffie-Hellman key exchange protocol in a relatively weak theory, perhaps in $\mathbf{VTC^0}$. For now, the author tries to prove correctness of the protocol in a certain extension of $\mathbf{VTC^0}$. This part of the thesis is quite badly written, so it is not that easy to follow the reasoning (for instance, it was unclear to me whether the notation $\mathsf{exp}_{G,P}$ stands for a new primitive or for a formula of some other language), but it is definitely clear that the proof of its main result contains highly problematic steps. For instance, the proof of Proposition 4.3, which is often used in the remainder of the section, is quite obviously incorrect. Later, at the end of the proof of the main result (pages 45-46), a certain formula is claimed to have 11 properties. Their proofs are so sketchy that it is often hard to tell what the argument is and whether it is correct, and to make matters worse, one of the less sketchy proofs (property no. 7) seems to use the so-called string induction scheme, which is not known (and not likely) to hold in $\mathbf{VTC^0}$. Needless to say, this does not inspire confidence in the proofs of the other properties.

Overall, therefore, the dissertation contains one (jointly-authored) new and interesting

result, a simple and elegant new proof (based on a potentially fruitful method) of a slightly weaker version of a known theorem, and, regrettably, some claimed results whose proofs are incorrect or of doubtful correctness. Fortunately, these doubtful results do not influence the remainder of the work, and they should probably be ignored as far as possible.

Presentation. The quality of the presentation in the thesis varies widely. In general, the appendices, both the technical Appendices A, B and Appendix C, which presents some background, are written tolerably well or better. Among these, Appendix A, with a generally careful exposition of a relatively complex proof, stands out, though even here one could point to some notational inconsistencies, not fully satisfying explanations, etc.

The main body of the thesis (Sections 1-5) deserves a more severe judgment. With the exception of Section 4, it does meet the "bare mininum" standard, in that it is possible to tell what the main results are and how they are (in outline) proved. However, overall this part of the dissertation is written rather carelessly and seems to have been prepared in haste. Typos are frequent, also in mathematical formulas; notation or terminology sometimes appears without a definition (an egregious case is Theorem 1.6, which additionally provides an incorrect reference); explanations are at times vague; some paragraphs should appear in a different order; and just before Theorem 1.27 there is even a paragraph which seems to have lost its second half. It is hard to avoid the feeling that many if not most of these lapses, which have a rather strong negative effect on readability, could have been eliminated with an additional two weeks' worth of effort.

Conclusion. It is apparent from the above discussion that S. Müller's dissertation does have some fairly significant deficiencies. In particular, the thesis would have been considerably improved by cleaning up the presentation, mentioning the relationship of the arithmetic part of Section 3 to Zambella's work, and most importantly, by not including Sections 3.2 and 4.

Nevertheless, the thesis does have its strong points. The most obvious of these is the main theorem of Section 2, which is a very nice result in propositional proof complexity and has already been accepted to the proceedings of the highly respected and competitive Logic in Computer Science conference. The work of Section 3, while not yet leading to a genuinely new result, does provide a new proof method with potential for applications. On the basis of these better, and more important, aspects of the dissertation, I am able to conclude that the author has demonstrated a very good understanding of the correspondence between arithmetic and propositional logic, some expertise in formalizing arguments in weak theories of arithmetic, and a general ability to perform creative mathematical research. The dissertation itself does on balance satisfy all the conditions normally required of a PhD thesis.

Leszek Kołodziejczyk

Appendix
More significant detailed issues not mentioned in the report

Main text:

Page 9, Theorem 1.5: Should not the "1.5" be "0.5"?

Top of page 10: it would be good to mention that the traditional definition of b.d. Frege disallows *all* formulas above a constant depth. The part just below this is written as if the author could not decide whether Theorem 1.8 or 1.9 should come first.

12_2: "proves"—viz., by means of short proofs.

13^{17}: "proofs are correct"—meaning, "sound"?

Top of page 15: some bounded theories are formulated in the thesis in such a way that it is not clear that they meet the definition of "bounded theory": cf. the induction axiom in the definition of $I\Delta_0$, and again the definition towards the bottom of page 17.

Page 15, Theorem 1.24: "Δ_0^b" should probably be "Δ_0". Neither has been defined.

16^2: the $|\cdot|$ function has not been introduced.

16^{10}: since we are talking about Ω_1, the lower index i should be 2 here.

16^{13-14}: "are equivalent". Only for strong enough theories; for example, *not* for $\mathbf{V^0}$.

17_{13-12}: "size" should be "length" twice.

Page 19, definition of $numones(X)$: why is there a "-1" here?

Page 20: the paragraph on "notation" should come before the one on "rearranging".

21_6: "their respective classes" is vague.

22_{11}: the "$\rightarrow$" here should be a "$\wedge$".

23^4: "the satisfaction relation"—for NC^1 circuits.

23^{19-18}: "$Z \models X$" is not Δ_1^B over $\mathbf{V^0}$! Cf. also Definition 1.45 and B.5.

24^{16}: the initial "$\forall \bar{x}$" should be deleted (cf. also Def. A.2).

24_2: the y, z superscripts next to m should be lowercase.

Page 26, Theorem 1.46: the first two parts of this theorem have provably false antecedents and consequents.

Page 31, main formula: the "$o(1)$" should be explained also *here*.

32^{13}ff.: should "$+o(1)$" be "$-o(1)$"? Just below that, I am not really sure how the expression for "satisfied once" is obtained.

Page 34, Corollary 3.5, "2^{n^δ}". Direct application of Theorems 3.7 and B.9 gives $2^{O(n^\delta)}$. Some fiddling with δ's and the $O(\cdot)$ notation gives $O(2^{n^\delta})$, but the $O(\cdot)$ seems needed.

Page 35, last paragraph of proof of Theorem 3.6: this argument is very sketchy.

Page 39: is T_1^2 here relativized or not? If relativized, then it is known that in general polylog cuts do not satisfy $\mathbf{VTC^0}$ (cf. Impagliazzo-Krajíček, MLQ 48(2002), 375-377).

39_6: "class" should be "theory" twice.

Bottom of page 41: how can a formula be the universal closure of a conjunction of (essentially) many copies of itself?

42^{13}: is MOD known to be total in $\mathbf{VTC^0}$?!

Page 44: a) In what theory do we have comprehension/induction for formulas with PROD? b) the "Commutativity" condition is incorrectly stated, c) one occurrence of "$\mathbf{V^0}$" should be "$\mathbf{VTC^0}$".

Appendices:

10^{5-6}: somewhere in the definition of imbalance, the word "difference" should appear.

Page 23, formula (A.10): I am not sure if this is fully correct.

Pages 36-37: the convention regarding ℓ in x_i^ℓ seems to change between page 36 and Definition A.44.

43^{4-3}: d is a bound on something else.

44_{11-14}: I found the explanations regarding "$o(1)$" confusing.

Page 50, Definition A.65 part 3.: should "ij" be "ji"?

55^{8}: "$(1/c')$" should be "$(1/n^{c'})$".

Page 58, Theorem A.75: It should be made explicit how the constants hidden in the $O(\cdot)$ notation relate to c from the definition of β. Otherwise, the "stems from direct computations" comment near the top of page 59 is unclear.

70_{16-14}: this way of translating boolean combinations may raise depth unnecessarily.

72_{6-5}, "whole sequence has length $O(n)$": which sequence, and length in what sense?

Page 73, Theorem B.13 and its proof: A itself, not just L, has to obey the time and space bounds. More importantly, there is some confusion here between m the logarithmically small element of N_1 and m as it appears in the time and space bounds. For instance, in the proof it seems that A is allowed to use n^ϵ space where n is the length of its input, but if $n = m^k$, and $\epsilon > 1/k$, then n^ϵ might not be logarithmically small, so configurations will not be coded by elements of N_1 and various quantifiers in the proof will not be first-order.

79^{11}: "image"?

86^{7-8}, "we will see this in Section C.1.2": we are in Section C.1.2.

Page 86, Theorem C.6: as stated, this is trivial.

Page 87, Theorem C.10: typical proofs of Gödel's second need a stronger base theory. Over very weak theories, the theorem becomes very dependent on the notion of consistency (e.g., in some cases holds for Hilbert-style provability but not known to hold for cut-free provability, etc.).

Page 88, Definition C.11 and below: the single-tape TM model is not really well-suited to complexity theory, especially for small resource bounds.

89_{15}: "space and time bound" is too vague to be informative, and TimeSpace is missing an argument.

89_{2}: the class of complements is not the complement of a class.

Pages 92-93: "L" normally stands for logarithmic space, but I expect that actually logarithmic *time* is meant here.