KTH Computer Science
and Communication

Current Research in Proof Complexity: Problem Set 4

Due: March 5, 2012. Submit as a PDF-file by e-mail to `jakobn at kth dot se` with the subject line `Problem set 4: ⟨your name⟩`. Solutions should be written in LaTeX or some other math-aware typesetting system. Please try to be precise and to the point in your solutions and refrain from vague statements. In addition to what is stated below, the general rules stated on the course webpage always apply.

Hints: For most or all problems, "hints" can be purchased at a cost of 5–10 points. In this way, you can configure yourself whether you want the problems to be more creative and open-ended, where sometimes a lot can depend on finding the right idea, or whether you want them to be more of guided exercises providing a useful work-out on the concepts of proof complexity. If you do not solve a problem, there is no charge for the hint (i.e., it is not deducted from the score on other problems).

Collaboration: Discussions of ideas in groups of two to three people are allowed—and indeed, encouraged—but you should write down your own solution individually and understand all aspects of it fully. For each problem, state at the beginning of your solution with whom you have been collaborating. Everybody collaborating on a certain problem is considered to have purchased a hint for that problem if one of the collaborators has done so.

Reference material: Some of the problems might be "classic" with solutions easily found on the Internet or in research papers. It is not allowed to use such solutions in any way unless explicitly stated otherwise. Anything said during the lectures on in the lecture notes should be fair game, though, unless you are specifically asked to show something that we claimed without proof in class. It is hard to pin down 100% formal rules on what all this means—when in doubt, ask the lecturer.

About the problems: Some of these problems are meant to be quite challenging and you are not necessarily expected to solve all of them. As a general guideline, a total score of around 100 points on this problem set should be enough to get a pass. Any corrections or clarifications will be posted on the course webpage `www.csc.kth.se/~jakobn/teaching/proofcplx11`.

1 (20 p) Let $G_{m,n}$ denote the grid graph with vertices (i,j) for $i \in [m]$, $j \in [n]$, and with edges from (i,j) to all $\{(k,\ell) \mid (k,\ell) \in \{(i-1,j),(i,j+1),(i+1,j),(i,j-1)\},\ k \in [m],\ \ell \in [n]\}$ (where we assume $m < n$). Let $G''_{m,n}$ be the multigraph having two copies of every edge in $G_{m,n}$. Consider the Tseitin contradiction $Ts(G''_{m,n}, f)$ for some odd-weight function f. Show that, as claimed in lecture 15, the formula $Ts(G''_{m,n}, f)$ can be refuted in resolution in length roughly $mn2^n$ and clause space roughly 2^n simultaneously. What are the best bounds you can get?

Hint: Consider the Tseitin formula as a set of mn (inconsistent) linear equations $E_{i,j} = \sigma_{i,j}$, one equation each for the local parity constraint around every vertex (i,j). Sum these equations one by one, row by row and columnwise in each row, to derive $0 = 1$. Show that resolution can simulate this linear algebra refutation with a blow-up that is not too bad.

2 (20 p) Suppose that ℓ_i and r_i, $i = 1, \ldots, k$, are natural numbers such that $r_i - \ell_i > 0$ and for all i we have $r_i - \ell_i > 2(r_{i-1} - \ell_{i-1})$. Prove that $\left|\bigcup_{i=1}^{k}\{\ell_i, r_i\}\right| \geq k + 1$ (as used crucially in the trade-off result by Beame, Beck and Impagliazzo covered in lectures 14 and 15).

Hint: One possible approach is to build a graph where the vertices are all distinct numbers in the set, and where the edges are (ℓ_i, r_i) for all i. Show that this graph is a tree, i.e., that it contains no cycles, and hence must have more vertices than edges.

3 (30 p) In lectures 7 and 8, we studied the result by Alekhnovich and Razborov that if the bipartite graph $G(F)$ associated to a CNF formula F is a good expander, then F is a hard formula with respect to PCR proof size (and thus also with respect to PC proof size). Is it possible to use an argument along these lines to prove that Tseitin contradictions over (constant-degree) expanders are also hard for PC/PCR with respect to size? For full credit, prove a strong lower or upper bound on Tseitin contradictions on expander graphs (via Alekhnovich–Razborov techniques or by other means).

4 (20 p) As defined in lecture 13, a Boolean function $f : \{0,1\}^d \mapsto \{0,1\}$ is *non-authoritarian* if fixing any variable x_i in $f(x_1, \ldots, x_d)$ to any value can never fix the truth value of f.

4a Can you give an example of a non-authoritarian function over two variables (i.e., of arity $d = 2$) other than exclusive or or the negation of exclusive or?

4b Can you give an example of a non-authoritarian function of arity $d = 3$ other than exclusive or or the negation of exclusive or over three variables or over a subset of two of the three variables?

Please do not forget to give a brief but convincing argument why any chosen functions are indeed non-authoritarian.

5 (20 p) With notation as in lecture 13, let $f : \{0,1\}^d \mapsto \{0,1\}$ be any non-constant Boolean function (not necessarily non-authoritarian), let $\mathbb{D}$ be a set of disjunctive clauses over $\mathit{Vars}^d(V)$, and let $V^* = \mathit{Vars}\big(\mathit{Rproj}_f(\mathbb{D})\big) \subseteq V$. Consider the bipartite graph with the vertices on the left labelled by clauses $D \in \mathbb{D}$ and the vertices on the right labelled by variables $x \in V^*$, and with edges between $D \in \mathbb{D}$ and $x \in V^*$ if some variable x_i appears in D. Prove that $N(\mathbb{D}) = V^*$, i.e., that all $x \in V^*$ have incoming edges from $\mathbb{D}$. (This was claimed without proof when we established that Rproj_f is a space-faithful projection with respect to resolution for the right f.)

6 (40 p) In lecture 12, we said that a function proj_f mapping $\mathcal{P}$-configurations to clause configurations is an f-projection if it is complete, nontrivial, monotone, and incrementally sound, and then proved a lemma that applying proj_f to any $\mathcal{P}$-refutation of the substituted formula $F[f]$ yields a resolution refutation of F. Give examples for all of the four conditions on projections explaining why this lemma fails to hold if the condition in question is omitted.

7 (50 p) In lecture 13, we claimed that $Sp_{\mathcal{R}}(F[\oplus] \vdash \bot) \geq \mathit{VarSp}_{\mathcal{R}}(F \vdash \bot)$ but only proved the weaker bound $Sp_{\mathcal{R}}(F[\oplus] \vdash \bot) = \Omega\big(\mathit{VarSp}_{\mathcal{R}}(F \vdash \bot)\big)$. The purpose of this problem is to understand the proof of the latter bound and improve the analysis to get the former bound.

7a Since we in fact did prove that $Sp(\mathbb{D}) \geq \mathit{VarSp}(\mathit{Rproj}_f(\mathbb{D}))$, a natural question (which was asked after the lecture) is why this is not sufficient to make Rproj_f exactly space-faithful. Explain why we can only claim that Rproj_f is linearly space-faithful.

7b If proj_f is an f-projection, then let its *monotone version* be $\mathit{proj}^*_f(\mathbb{D}) = \bigcup_{\mathbb{D}' \subseteq \mathbb{D}} \mathit{proj}_f(\mathbb{D}')$. Prove that the monotone version proj^*_f of any projection is also (as the terminology suggests) a projection in the sense defined in lecture 12.

7c Prove that the monotone version Rproj^*_f of Rproj_f is exactly space-faithful and that this implies $Sp_{\mathcal{R}}(F[\oplus] \vdash \bot) \geq \mathit{VarSp}_{\mathcal{R}}(F \vdash \bot)$.