Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Institute of Theoretical Computer Science
Emo Welzl, lecturer, and Robin Moser, teaching assistant

Tuesday, May 8, 2012

Satisfiability of Boolean Formulas Spring 2012
Special Assignment Set 3

- The solution is due on **Friday, May 25, 2012**. Please bring a print-out of your solution with you to the lecture. If you cannot attend (and please only then), you may alternatively send your solution as a PDF to `robin.moser@inf.ethz.ch`. We will send out a confirmation that we have received your file. Make sure you receive this confirmation within the day of the due date, otherwise complain timely.
- Please solve the exercises carefully and then write a nice and complete exposition of your solution using a computer, where we strongly recommend to use LaTeX. A tutorial can be found at `http://www.cadmo.ethz.ch/education/thesis/latex`.
- For geometric drawings that can easily be integrated into LaTeX documents, we recommend the drawing editor IPE, retrievable at `http://ipe7.sourceforge.net/` in source code and as an executable for Windows.
- You are welcome to discuss the tasks with your colleagues, but we expect each of you to hand in your own, individual write-up.
- There will be three special assignments this semester. Each of them will be graded and the average grade will contribute 30% to your final grade.
- This is a theory course, which means: if an exercise does not explicitly say "you do not need to prove your answer" or "justify intuitively", then a formal proof is **always** required.
- As with all exercises, the material covered in special assignments is relevant for the final exam.

Exercise 1 (Mixed Tasks on PPZ / PPSZ) *(35 Points)*

(a) Consider the 3-SAT version of the formula $F_n^{\sim}$ from the last special assignment: a formula over $\{x_1, x_2, \ldots, x_n\}$ with all 7 clauses over $\{x_1, x_2, x_3\}$ containing at least one positive literal as well as all clauses $\{\bar{x}_{i-2}, \bar{x}_{i-1}, x_i\}$ for $4 \leq i \leq n$. Prove that PPZ needs (expected) exponential time to solve this formula and that on the other hand PPSZ for $D = \log n$ solves the formula in (expected) subexponential time.

HINT: You may want to use a Chernoff Bound at some point.

(b) Let F be a $(\leq k)$-CNF formula F over variables V. Let us call a satisfying assignment $\alpha \in \mathrm{sat}_V(F)$ *r-superisolated* if for every variable $x \in V$, $F^{[x \mapsto (1-\alpha(x))]}$ contains at least r independent clauses violated by α. Prove that PPZ has subexponential success probability on formulas that have a $(\log n)$-superisolated assignment.

(c) Given a k-uniform hypergraph H over n vertices V, a *proper k-coloring* is a coloring $\gamma : V \to \{1..k\}$ of the vertices V with k colors such that no edge becomes monochromatic. We are interested in the computational problem of deciding whether a given hypergraph is properly k-colorable.

Consider the following PPZ-like approach to the problem: process the vertices in a uniformly random ordering and choose for each vertex a uniformly random color among those colors that do not lead to closing a monochromatic edge. Assume that H admits an n-isolated coloring, i.e. a coloring for which, if you change any single color, the coloring is no longer proper. Determine a

bound[1] depending on k for the success probability of this algorithm in finding a proper coloring.

Compare[2] this approach to the method of random downsampling, where we randomly forbid $k-2$ of k colors at each vertex and then run Schöning's algorithm (or PPSZ) on the resulting k-CNF formula. Which is the method of choice for which k?

Exercise 2 (PPZ with Fewer Random Bits) *(35 Points)*

(a) Prove: Let k be some constant. There exists a code $\mathcal{C} \subseteq \{0,1\}^r$ of length $r = \mathcal{O}(\log n)$ of size $|\mathcal{C}| = \theta(\sqrt{n})$ with the property that for any k-tuple $c_1, c_2, \ldots, c_k \in \mathcal{C}$ of codewords, there exist indices $i_1, i_2, \ldots, i_k$ such that c_j is (among the k-tuple) the unique codeword where the i_j-th bit is one, for all $1 \leq j \leq k$. Moreover, this code can be generated deterministically in subexponential time.

(b) Use (a) to prove: Let k be some constant. There exists a code $\mathcal{C}' \subseteq \{0,1\}^r$ of length $r = \mathcal{O}(\log^2 n)$ and of size $|\mathcal{C}'| = n$ with the following property. Let $Y_1, Y_2, \ldots Y_r$ be a supply of mutually independent random variables distributed uniformly among the numbers $\{0..n-1\}$. Let $\{c_1, c_2, \ldots, c_n\} = \mathcal{C}'$ be the codewords. Consider the n random variables $X_1, \ldots, X_n$ defined as

$$X_i := \sum_{j=1}^{r} (c_i)_j \cdot Y_j \pmod{n}.$$

Then for any k-tuple $i_1, i_2, \ldots, i_k \in \{1..n\}$ of pairwise distinct indices, the probability that X_{i_1} takes the unique maximum value among $X_{i_1}, \ldots, X_{i_k}$ equals $1/k + o(1)$. Moreover, such a code can be generated deterministically in subexponential time.

(c) Use (b) to modify the PPZ algorithm in such a way, that called on k-CNF formulas F which are guaranteed to have an n-isolated satisfying assignment, the algorithm runs in subexponential time, outputs a satisfying assignment with probability at least $2^{-(1-\frac{1}{k})n+o(n)}$, and consumes a total of $(1-1/k)n + o(n)$ random bits.

Exercise 3 (Typical Executions) *(30 Points)*

Consider the Markov chain $\{D_i \in \mathbf{Z}\}_{i \in \mathbf{N}}$ we have used to analyse Schöning's algorithm for 3-SAT, i.e. $D_0 \sim \mathrm{Bin}(n, \frac{1}{2})$ and then for all $1 \leq i \leq 3n$, $D_i = D_{i-1} - 1$ with probability $1/3$ while $D_i = D_{i-1} + 1$ with probability $2/3$. Recall that we have called an event E *typical* if $\Pr[E \wedge \exists i \leq 3n : D_i = 0] = (3/4)^{n+o(n)}$.

For each of the following events (separately), prove or disprove that the event is typical.

(a) $E_1 = \{\forall i \leq \min\{j \mid D_j = 0\} : D_i \leq D_0\}$.

(b) $E_2 = \{\forall i \leq \min\{j \mid D_j = 0\} : \lfloor \frac{D_i}{100} \rfloor \leq \lfloor \frac{D_{i-1}}{100} \rfloor\}$.

(c) $E_3 = \{|\{i \leq \min\{j \mid D_j = 0\} \mid D_i = D_{i-1} - 1\}| \equiv_2 0\}$.

[1] you can give this bound as a formula still involving integrals, you need not evaluate them

[2] you can do this numerically; use a computer algebra program to evaluate your formula for a number of values until you see a tendency