Topics in Algorithms- Random SAT (0510-7410) - Spring 2014
Benny Applebaum

Please submit a printed copy of your solution by May 13.

Problem Set #2

Question 1: Pure Literal

Generalize the analysis of the pure literal heuristic to the case of random k-CNF formulas. Recall that in this distribution, denoted by $F_k(n, m)$, each of the m clauses is chosen independently at random among all $2^k\binom{n}{k}$ possible k-clauses. What is the largest clause/variable density $r_k = m/n$ for which the algorithm succeeds with constant probability?

Question 2: Expansion

An (n, m, d) graph is a bipartite graph with n left nodes L (variables), and m right nodes R (constraints) where the degree of each right node is d. Let $S \subseteq R$ be a set of right nodes. A left node is a *neighbor* of S if it is connected to one or more nodes in S. A left node is called a *unique neighbor* of S if it is connected to a single node in S. We let $N(S)$ denote the set of neighbors of S, and let $U(S)$ denote the set of unique neighbors of S. An (n, m, d) graph is called (k, α) expander if every set $S \subseteq R$ of size at most k has at least $\alpha|S|$ neighbors, namely, $|N(S)| \geq \alpha|S|$. The graph is (k, α) *unique expander* if every set $S \subseteq R$ of size at most k has at least $\alpha|S|$ unique neighbors, namely, $|U(S)| \geq \alpha|S|$.

1. Prove that (k, α) expansion is impossible for $k > n/\alpha$. Similarly, show that (k, α) expansion is impossible for $\alpha > d$.

2. Let $\varepsilon > 0$ be an arbitrarily small constant. Prove that a *random* $(n, m, d = O(\frac{\log(m/n)}{\varepsilon}))$ graph is, whp, an (k, α) expander where $k = \Theta(\frac{\varepsilon n}{d})$ and $\alpha = (1-\varepsilon)d$.
   Note: A *random* (n, m, d) graph is selected by connecting each right node independently to d random distinct left nodes.

3. Let $\varepsilon < 1/2$. Show that if an (n, m, d) graph G is $(k, (1-\varepsilon)d)$ expander then it is also $(k, (1-2\varepsilon)d)$ unique expander.

Question 3: Random Strings cannot be compressed

Let $f : \{0,1\}^n \to \{0,1\}^*$ be an injective function. We say that f compresses an input x by b bits if the length of $f(x)$ is at most $n - b$. It is not hard to see that f cannot compress all the inputs (due to the pigeonhole principle). In this question we will prove a stronger version of this claim. Let $\delta > 0$ be a real number. Show that there are no more than $\delta 2^n$ strings $x \in \{0,1\}^n$ which f compresses by $\lg(\frac{1}{\delta})$ bits.

Question 4: Random Walk

We saw that below the pure-literal threshold, i.e., $r < 1.63$ the Random Walk algorithm terminates whp in polynomial time. In this question, we will see that the running time in this case is actually *linear*. **Definitions**: Let A be a 3-CNF over n variables where we view each clause a as a vector in $\{-1, 0, -1\}^n$ by setting the i-th coordinate of a to be 1 if x_i appears positively in a, -1 if x_i appears negatively in a, and 0 if x_i does not appear in a. Recall that a weighted assignment $\beta \in \mathbb{R}^n$ is a *terminator* for 3-CNF formula A if for every clause a of A we have $\langle \beta, a \rangle = \sum_i \beta_i \cdot a_i \geq 1$. In the following $m/n < 1.63$.

1. Suppose that A has the following property: if we put 4 tokens on each variable, then it is possible to distribute them among the clauses such that each clause takes exactly 7 tokens, where a clause can take a tokens only from its members. Show that A has a terminator β with $|\beta_i| \leq 4$ for all i's.

2. We say that A is $7/4$-expanding if every subset S of clauses "touches" at least $7/4|S|$ variables. Prove: if A is $7/4$-expanding then it satisfies the property of the previous sub-question. **Hint**: use Hall's theorem.

3. Let A be a random formula chosen from $F_3(n, m)$, and let A_d be the resulting formula after d iterations of the pure-literal heuristic. Argue that there exists a constant d such that, whp, A_d is $7/4$-expanding. **Hint**: By using the lemma proven in class it suffices to show that for some constant α a random formula from $F_3(n, m)$ is, whp, $(\alpha m, 7/4)$-expanding, namely: every subset S of at most αm clauses "touches" at least $7|S|/4$ variables.

4. Show that, whp, $A \stackrel{R}{\leftarrow} F_3(n, m)$ has a terminator β with $|\beta_i| < O(1)$ for all i's. Conclude, (by using the lemma proven in class) that Random-Walk terminates in linear-time. **Hint:** Partition A to two 3-CNFs: A_1 which is likely to be satisfied after $d = O(1)$ iterations of pure-literal, and A_2 which is $7/4$-expanding. Carefully combine the terminator of A_1 (shown in class) with terminator of A_2.

Question 5: Implementation

Implement Pure Literal and Random Walk. Submit the code and the following graphs:

1. For $r_0 = 1.6$ generate 20 instances of $F_3(n, rn)$ for each of the following values $n = 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000$. Plot a grid whose x-axis is the number of variables and the y-axis is the number of steps performed by the algorithm, and put a blue point in the grid for every trial.

2. Conduct the same experiment with $r_1 = 1.7$. Add the points that represent this experiment to the previous grid as red points.

3. Repeat the previous questions with Random Walk and $r_0 = 2, r_1 = 3$.