ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 4, Issue 2, pp: (259-261), Month: April - June 2016, Available at: www.researchpublish.com

AN EFFICIENT VERSION OF SCHONING'S ALGORITHM APPLIED FOR ONE-IN-THREE SAT

Cristian Dumitrescu

Independent Mathematician, Kitchener, Canada

Abstract: **In this article I describe an efficient, randomized algorithm that finds a solution to the ONE-IN-THREE SAT problem (when one exists) in polynomial time with high probability.**

Keywords: **The Satisfiability Problem, ONE-IN-THREE SAT, Markov chain, random walk with absorbing barriers, NP-complete.**

1. INTRODUCTION

We will consider the ONE-IN-THREE SAT problem, a variant of 3SAT for which every clause has **exactly** one literal in the clause true (instead of at least one). In [1] it is proved that ONE-IN-THREE SAT is NP-complete.

Proposition 1. ONE-IN-THREE SAT is NP-complete.

Proof. For the proof see [1], chapter 9, page 207, problem 9.5.3.

Lemma 1 (see [1], lemma 11.2, page 247)**.** If X is a random variable taking nonnegative integer values, then for any $k > 0$ the probability $\text{prob}\{X \geq k \cdot E(X)\} \leq \frac{1}{k}$, where E(X) denotes the expected value of X.

Proof. Let p_i be the probability that $= i$. Then we have:

$E(X) = \sum_i i \cdot p_i = \sum_{i \leq k \cdot E(X)} i \cdot p_i + \sum_{i > k \cdot E(X)} i \cdot p_i > k \cdot E(X) \cdot \text{prob}\{X > k \cdot E(X)\}$. From the relation above the lemma follows.

In this article I will present an efficient probabilistic algorithm that solves ONE-IN-THREE SAT in polynomial time with high probability.

2. THE PRESENTATION OF THE ALGORITHM

Definition 1. Given a truth assignment for the n variables in a ONE-IN-THREE SAT instance, the **dual assignment** is simply obtained by flipping the truth value for each variable (from 0 to 1 and from 1 to 0).

Definition 2. Given a truth assignment for the n variables in a ONE-IN-THREE SAT instance, we have two types of flips, the $0 \rightarrow 1$ flip (when we change the corresponding variable so that a literal changes its truth value from 0 to 1), and the $1 \rightarrow 0$ flip (when we change the corresponding variable so that a literal changes its truth value from 1 to 0).

We are now ready to present the algorithm. This algorithm is a modified version of Schoning's algorithm.

Modified Schoning Algorithm.

Input: a formula in ONE-IN-THREE SAT 3-CNF with n variables.

Guess an initial assignment $a \in \{0,1\}^n$ ***for the n variables, uniform at random.***

Repeat $C \cdot n^2$ ***times:***

ISSN 2348-1196 (print)
International Journal of Computer Science and Information Technology Research ISSN 2348-120X (online)
Vol. 4, Issue 2, pp: (259-261), Month: April - June 2016, Available at: www.researchpublish.com

Randomly choose one type of flip to perform.

If the formula is satisfied by the dual assignment: stop and accept.

If the formula is satisfied by the actual assignment: stop and accept.

If there are no literals left that allow the type of flip chosen: stop and reject.

Choose one literal at random (from any clause of the expression) that allows the type of flip chosen, and flip its truth value.

There are a few differences between this algorithm and Schoning's algorithm (see [2]) and the present algorithm. We focus on the overall number of possible flips, not just on one clause at a time. This is, in fact, a generalization of Schoning's algorithm. In the next section I will outline a proof that in a sufficiently large number of steps that does not exceed a quadratic polynomial of n, the algorithm will find a satisfying solution (if it exists) with high probability.

3. ANALYSIS OF THE ALGORITHM

We consider a ONE-IN-THREE SAT problem with n variables. We define as usual the Hamming distance from an actual truth assignment for the n variables to a satisfying assignment (if it exists, and if there are more than one just choose one). The Hamming distance is the number of variables that have to be flipped in order to get to the satisfying assignment. This Hamming distance will be contained in the set $\{0, 1, 2, \ldots \ldots . n\}$, and will also represent the state of the associated Markov chain (this method has also been used by Schoning). In fact, we will be dealing with one dimensional random walks with two absorbing barriers, but we may also refer to them as Markov chains..

We are ready to state the theorem.

Theorem 1. For all (including the hardest) ONE-IN-THREE SAT problems, the algorithm presented in section 2 will find a satisfying assignment or its dual (finding its dual is equivalent to finding a satisfying assignment) in polynomial time (quadratic time) with high probability. For sufficiently large constant C (and by running the algorithms several times), we can make the probability that the algorithm will fail (meaning that it will report that the expression is unsatisfiable when in fact it is) with probability arbitrarily small.

Proof. We have two types of flips that appear in the algorithm.

The $0 \rightarrow 1$ flip.

In this case the Hamming distance increases by 1 with probability $\frac{2}{3}$.

In this case the Hamming distance decreases by 1 with probability $\frac{1}{3}$.

The $1 \rightarrow 0$ flip.

In this case the Hamming distance increases by 1 with probability $\frac{1}{3}$.

In this case the Hamming distance decreases by 1 with probability $\frac{2}{3}$.

This happens because in the context of ONE-IN-THREE SAT, a satisfying clause has exactly one literal with the truth value 1 and two literals with the truth value 0.

We note that in the algorithm, both the $0 \rightarrow 1$ flips and the $1 \rightarrow 0$ flips are chosen with probability $\frac{1}{2}$. That means that overall, in the inner loop of the algorithm, the Hamming distance will increase with probability $\frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2}$. Also we see that overall, the Hamming distance will decrease also with probability $\frac{1}{2} \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{1}{2}$. We have here a symmetric (unbiased) random walk with two absorbing barriers, when the Hamming distance is 0 (when the satisfying assignment is found), or at Hamming distance n (when the dual of a satisfying assignment is found).

We know that for a one dimensional, perfectly symmetric random walk with two absorbing barriers (and with n states), the expected lifetime is quadratic in n, and that is the most symmetric, balanced case (and this is a well known result, I will not give the proof here). From lemma 1 we see that the probability that the absorption time (seen a s a random variable) is higher than quadratic (when a solution exists) is vanishingly small.

The problem can also be formulated in terms of the expected duration of the game for a version of the gambler ruin problem (for an unbiased game), but I will not emphasize that here.

In any case, we have quadratic time until absorption with arbitrarily high probability.

Let's assume that we start with a 3SAT problem with n' variables and m' clauses (we emphasize this case because SAT problems are most studied). We write $' = \alpha \cdot n'$, where α is a small positive constant. For the most difficult problems α is around 4.267. The equivalent ONE-IN-THREE SAT problem will have $n' + 4 \cdot m'$ variables and $3 \cdot m'$ clauses (I assume that the reader is familiar with the reduction of a 3SAT problem into ONE-IN-THREE SAT, this is a well known procedure). In other words, the equivalent ONE-IN-THREE SAT problem will have $(4 \cdot \alpha + 1) \cdot n'$ variables and $3 \cdot \alpha \cdot n'$ clauses (we understand that we take the integer part of these expressions, but for simplicity we will omit that). That means that the constant C must be greater than $(4 \cdot \alpha + 1)^2$ (for the most difficult problems). We note that the optimum value of the constant (for a given 3SAT instance) depends on ratio between the number of clauses and the number of variables of the expression.

By taking the constant C large enough in the algorithm, and if we run the algorithm several times, we can make the probability of failure as small as we want. QED.

Observation: We must be careful and see what happens when there are no literals left that allow the chosen type of flip to be performed, for example if all literals have the truth value 0 (or all 1) in the current assignment. That is not likely to happen for the interesting problems where we have both positive literals (the associated variable itself) and negative literals (the negation of the associated variable). Also this situation (or any possible infinite deterministic loops) happen with vanishing probability, and by running the algorithm many times we can solve this problem.

4. DISCUSSION AND CONCLUSIONS

For general implications, related to efficiently solving NP – complete problems, see [3]. An interesting application is related to the problem of automated theorem proving using an efficient algorithm for NP – complete problems. The impact of this type of algorithm in mathematics, cryptography, science in general is hard to estimate. Also note that a derandomization of these algorithms could lead to interesting results. In a few words, this is a very important results with extremely important applications in any field of activity.

ACKNOWLEDGEMENT

Credit must be given where credit is due. Without Schoning's algorithm this type of algorithm would not exist. Schoning's algorithm hits the satisfying assignment in at most exponential time (in the number of variables). The trick here is to take into consideration the dual assignment, focus on the ONE-IN-THREE SAT problem, and focus on the two types of flips possible. I also emphasize that Professor Fortnow's book (reference [3]) is a great source of inspiration and motivation that led me to seriously consider this problem. Also I emphasize that both Professor Schoning and Professor Fortnow corrected some errors that I had in some previous versions of this paper, and for that I am very grateful, but that does not mean that they endorse this version of my paper.

REFERENCES

[1] C.H. Papadimitriou, *"Computational Complexity "*, Addison - Wesley Publishing Company, Inc., 1994.

[2] Uwe Schoning, *"A probabilistic Algorithm for k-SAT and Constraint Satifaction Problems"*, Research Supported by the ESPRIT Basic Research, 1991.

[3] L. Fortnow, *"The Golden Ticket, P, NP, and the Search for the Impossible "*, Princeton University Press, 2013.

AUTHOR'S PROFILE:

Cristian Dumitrescu. Biographical notes. I was born in 1964 in Romania. In 1988 I graduated the University of Bucharest, Faculty of Mathematics with a BSc. In Mathematics. I worked as a high school teacher of mathematics and computer programmer for various companies and software houses in Romania, UK and Canada. I have been a Canadian citizen since 1998. In the last 20 years I have been working in a different field of activity, but I have always kept close to my heart mathematics and physics.