Global Journal of Computer Science and Technology
Neural & Artificial Intelligence
Volume 13 Issue 1 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 0975-4172 & Print ISSN: 0975-4350

A Proposed SAT Algorithm

By Bagais A., Junaidu S. B. & Abdullahi M.
Ahmadu Bello University, Zaria-Nigeria

Abstract - This paper reviews existing SAT algorithms and proposes a new algorithm that solves the SAT problem. The proposed algorithm differs from existing algorithms in several aspects. First, the proposed algorithm does not do any backtracking during the searching process that usually consumes significant time as it is the case with other algorithms. Secondly, the searching process in the proposed algorithm is simple, easy to implement, and each step is determined instantly unlike other algorithms where decisions are made based on some heuristics or random decisions. For clauses with three literals, the upper bound for the proposed algorithm is $O(1.8171^n)$. While some researchers reported better upper bounds than this, those upper bounds depend on the nature of the clauses while our upper bound is independent of the nature of the propositional formula.

Keywords : *propositional satisfiability, NP-complete, complexity, complete algorithms.*

GJCST-D Classification : *F.2.1*

Strictly as per the compliance and regulations of:

A Proposed SAT Algorithm

Bagais A.[α], Junaidu S. B.[σ] & Abdullahi M.[ρ]

Abstract - This paper reviews existing SAT algorithms and proposes a new algorithm that solves the SAT problem. The proposed algorithm differs from existing algorithms in several aspects. First, the proposed algorithm does not do any backtracking during the searching process that usually consumes significant time as it is the case with other algorithms. Secondly, the searching process in the proposed algorithm is simple, easy to implement, and each step is determined instantly unlike other algorithms where decisions are made based on some heuristics or random decisions. For clauses with three literals, the upper bound for the proposed algorithm is $O(1.8171^n)$. While some researchers reported better upper bounds than this, those upper bounds depend on the nature of the clauses while our upper bound is independent of the nature of the propositional formula.

Keywords : propositional satisfiability, NP-complete, complexity, complete algorithms.

I. Introduction

Propositional satisfiability (SAT) is one of the classical problems in Computer Science. The importance of SAT comes from the fact that a large class of real-world problems can be expressed in terms of a SAT instance and that it was the first problem proven to be NP-Complete (Cook, 1971). The SAT problem has a wide range of practical real world applications (Barbour, 1992; Crawford & Baker, 1994; Devadas, 1989; Kauts & Selman, 1992; Larrabee, 1992). Many algorithms, categorized into complete and incomplete algorithms, were proposed to solve this problem efficiently over the last decades.

Complete algorithms can state whether a SAT instance is satisfiable giving the satisfying assignments or unsatisfiable giving a 'no' answer. Incomplete algorithms can only give an answer of 'yes' for satisfiable SAT instances only but cannot give an answer for unsatisfiable instances.

This paper proposes a new complete algorithm that differs from the ones in the literature in the following aspects:

- No backtracking during the searching process that usually consumes significant amount of time.
- Has a simple, deterministic and easy to implement search process, unlike other algorithms where decisions are either made randomly or based on some heuristics.

The remainder of the paper is structured as follows. Section 2 describes the proposed algorithm with the aid of an example. Section 3 captures the algorithm in pseudo code while Section 4 presents the complexity analysis of the algorithm. We present related work in Section 5. Sections 6 and 7 summarize and provide references, respectively.

II. Illustrating the Proposed Algorithm

Unlike other algorithms that make a decision on a single value (true/false) for a variable x, the proposed algorithms takes into consideration all satisfying assignments for a clause C and use them for the next clauses so that backtracking is avoided.

Consider the following formula: $F = (x_1 \vee x_3 \vee \bar{x}_4) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_5) \wedge (x_2 \vee \bar{x}_3 \vee x_4)$

The first clause can be satisfied by any of the following assignments $x_1 = true, x_3 = true, x_4 = false.$ The algorithm tries to find assignments for all variables in clause while preserving at least one of the given assignments for x_1, x_3, or x_4 in the first clause.

In general, the process starts from the first clause c_1 and produces the set of assignments that satisfy c_1 which obviously are the literals in that clause.

If the clause has *k* literals, then *k* assignments can satisfy it (as in the previous formula, the first clause has three assignments). In the next step, the set of assignments that satisfy the set of previous clause(s) are checked with all the literals of the next clause. The process continues until all the clauses in the formula are covered, after which the resulting set of assignments each satisfies the formula.

When a set of assignments from previous clause(s) is checked with the literals of the current clause, each literal may *agree*, *disagree* or be *neutral* to the assignment. A literal agrees with an assignment when the assignment includes the literal. A literal disagrees with an assignment when the assignment includes a negation of the literal. A literal is neutral to an

Author α : Department Information and Computer Science, King Fahd University of Petroleum and Minerals, Dahram-Saudi Arabia. E-mail : bagais2008@gmail.com
Author σ ρ : Department of Mathematics, Ahmadu Bello University, Zaria-Nigeria. E-mails : abuyusra@gmail.com, muham08@gmail.com

assignment when the assignment neither agrees nor disagrees with the literal.

$$F = (x_1 \vee x_3 \vee \bar{x}_4) \wedge (\bar{x}_1 \vee x_2 \vee \bar{x}_5) \wedge (x_2 \vee \bar{x}_3 \vee x_4)$$

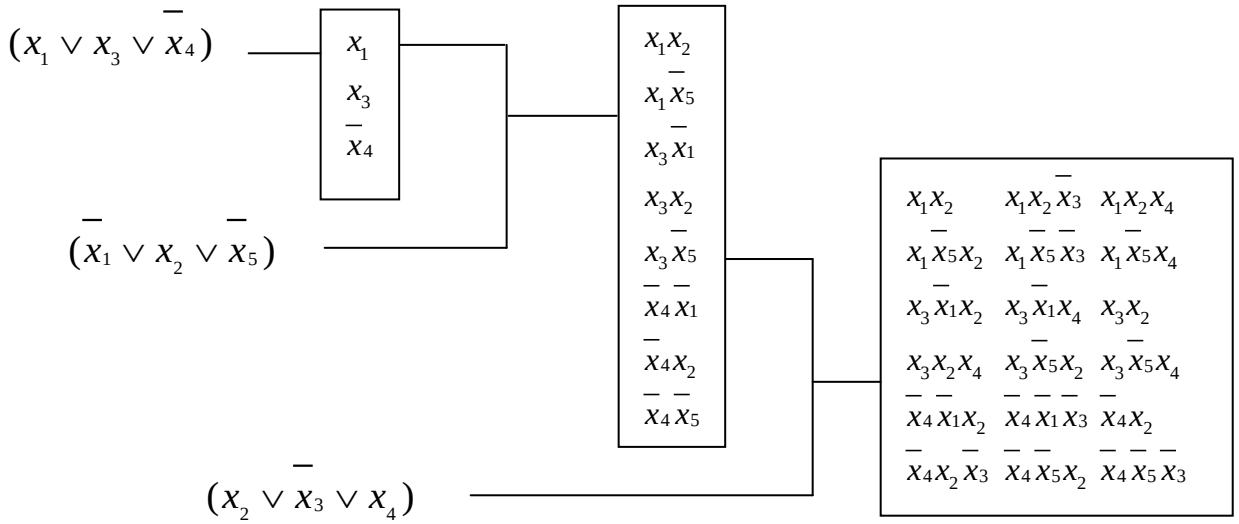

Figure 1 : Assignment Production

In the first step, the satisfying assignments for the first clause are its literals. The assignments produced for the first clause are shown in the top-left rectangle in *Figure 1*. Each of these assignments is checked with the literals of the second clause, $(\bar{x}_1 \vee x_2 \vee \bar{x}_5)$. The assignment of x_1 disagrees with the first literal of the second clause, $\bar{x}_1$ resulting in no assignment produced. The same assignment, x_1 is checked with the second literal, x_2. Since this literal is neutral to x_1, a new assignment is produced by combining x_1 and x_2, as shown in the middle rectangle in *Figure 1*. Next, x_1 is checked with $\bar{x}_5$, giving $x_1\bar{x}_5$, since $\bar{x}_5$ is neutral to x_1. Similarly, the assignments x_3 and $\bar{x}_4$ are checked with the literals of the second clause leading to six additional assignments as shown in the middle rectangle of *Figure 1*. To complete this example, the literals of the third clause are checked with these eight assignments producing the 18 new assignments in the right-most rectangle of *Figure 1*. Note that each of these 18 assignments satisfies the given formula.

Note that when an assignment agrees with the clause in consideration, the process might produce shorthand for $x_1 \vee x_2$ etc. We will illustrate this with the pair of clauses:

$$(x_1 \vee x_2 \vee x_3)$$
$$(x_1 \vee x_4 \vee x_5)$$

The satisfying assignments for this pair of clauses are:

$x_1x_1 (or\ x_1)$	x_2x_1	x_3x_1
x_1x_4	x_2x_4	x_3x_4
x_1x_5	x_2x_5	x_3x_5

From this group, it can be seen that the assignments $\{x_1x_4, x_1x_5, x_2x_1, x_3x_1\}$ are subsumed in the first assignment x_1. This is because each of these assignments produces the same result as x_1.

Thus, these assignments can be dropped to avoid redundancy. Therefore, *Figure 1* can now be redrawn without the subsumed assignments as shown in *Figure 2*.

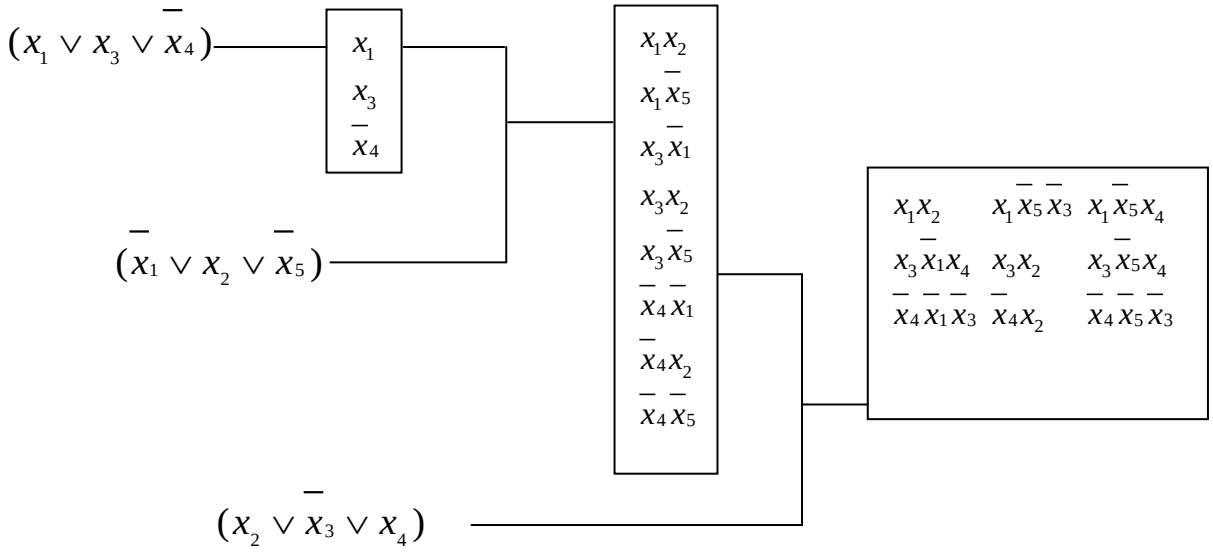

Figure 2 : Satisfying assignments without redundancies

Since the subsumed assignments are produced from clauses that have a literal in common, the proposed algorithm starts by extracting all clauses that do not share a literal. For a satisfiability formula with n literals each clause containing exactly k literals, the minimum number of clauses in which no two clauses have a common literal is $\frac{2n}{k}$.

$$F = (x_1 \vee x_2 \vee \bar{x}_4) \wedge (\bar{x}_1 \vee x_3 \vee x_4) \wedge (x_1 \vee \bar{x}_3 \vee x_6) \wedge (x_5 \vee \bar{x}_2 \vee \bar{x}_6) \wedge (x_4 \vee \bar{x}_5 \vee x_2) \wedge (x_3 \vee \bar{x}_6 \vee \bar{x}_5)$$

For example, we need at least 4 clauses to have the 12 literals in the following formula. But because of the distribution of literals, we need 5 for that purpose. Therefore, the algorithm will extract the clauses that do not have common literals. There are two advantages in doing so:

1. The algorithm will save the time to check the existence of subsumed assignments which is a process that consumes an amount of time equal to the number of assignments.
2. The time complexity of the algorithm becomes easier to prove (see Section 4).

Theorem 1

Consider a satisfiability formula with m clauses each of k literals. An agreement between an assignment and a literal in the i^{th} clause produces at least $\begin{cases} 2(k-1)k^{m-i}; \ i=2 \\ (k-1)k^{m-i}; \ 3 \le i \le m \end{cases}$ redundant assignments.

Proof: (By induction).

Base Case

The base case is when $i=m$ and the total number of redundant assignments will be $(k-1)k^{m-m} = (k-1)k^0 = (k-1)$. Clearly, the theorem holds for $i=m$.

Inductive Hypothesis

Suppose the theorem holds for $i=2,3,4,...,p$ for some clause $2 \le p < m$. The total redundant assignments will be $\begin{cases} 2(k-1)k^{m-p}; \ p=2 \\ (k-1)k^{m-p}; \ 3 \le p < m \end{cases}$. If a literal with which an assignment agrees is in $p+1$ clause, then the total redundant assignments will be

$$\begin{cases} \frac{2(k-1)k^{m-p}}{k} = 2(k-1)k^{m-p-1} = 2(k-1)k^{m-(p+1)}; \ i=2 \\ \frac{(k-1)k^{m-p}}{k} = (k-1)k^{m-p-1} = (k-1)k^{m-(p+1)}; \ 3 \le i \le m \end{cases}.$$

That is, the theorem holds for $p+1$. By induction on p, the theorem is true for all values of i.

Theorem 2

Consider a satisfiability formula with m clauses each of k literals. A disagreement between an assignment and a literal in the i^{th} clause reduces the number of assignments by at least by $k^{m-i}; \ 2 \le i \le m$.

Proof: (By induction)

Base Case

The base case is when $i=m$ and the total number of assignments will be reduced by $k^{m-m}=k^0=1$. Clearly, the theorem holds for $i=m$.

Inductive Hypothesis

Suppose the theorem holds for $i=2,3,4,...,p$ for some clause $2 \le p < m$. The total assignments will be reduced by k^{m-p}. If a literal with which an assignment agrees with is in $p+1$ clause, then the total assignments will be reduced by $\frac{k^{m-p}}{k}=k^{m-p-1}=k^{m-(p+1)}$. That is, the theorem holds for $p+1$. There by induction on p, the theorem is true for all values of i.

III. The Proposed Algorithm Pseudocode

The most important step in any complete or incomplete SAT algorithm is the decision over the value of a given variable in the formula. If the decision on that variable is wrong, the algorithm will waste its time searching for a solution before it discovers that the value assigned to the variable does not lead to a satisfying assignment and consequently a backtrack is done to change that value. The problem with making a decision for a variable x using the heuristics is that they do not consider how this decision or assignment will affect other related variables that appear in the same clauses as the variable x. If the search process keeps all possible assignments that satisfy a clause before moving forward, then no backtrack is needed. Instead, these assignments can be used to determine the values of variables that satisfy the next clauses. In the case that none of the variables in the current clause agrees with all the assignments, then the formula is unsatisfiable. This leads to the main idea of the proposed algorithm for assigning values to the variables.

The Algorithm

Input: *F[m]; //formula with m clauses*
Output : *A[k^m]; //Possible assignment satisfying m clauses.*

```
1.  getDistinctClauses(F[m]);

2.  For i = 1 to disticntclauses.length – 1;//number of distinct clauses
For j = 1 to k //k is the number of literals in a clause
LIT[i][j] := disticntclauses[i];
        End for
    End for

3.  For i = 1 to k
A[i] := LIT[1][i]; //literals of the first clause(initial set of satisfying substitutions)
End for

4.  For i = 2 to disticntclauses.length;//number of distinct clauses
For j = 1 to k
generateAssignment(LIT[i][j], A[], temp[]);
//A[] contains the set of satisfying substitutions from previous clauses
//temp[] contains assignments formed by combining assignments in A[]  with a literal LIT[i][j]
End for
A[] := A[] + temp[];
  End for

5.  For i = distinctclauses.length + 1 to m;//number of distinct clauses
For j = 1 to k //k is the number of literals in a clause
LIT[i][j] := nondistinctclauses[i];
        End for
    End for

For i = distinctclauses.length to m
For j=1 to k
```

```
generateAssignment(LIT[i][j], A[], temp);
End for
removeSubsumedAssignments(tempassignments[], arraysubsumed[]);
A[] := A[] + temp;
    End for

6.   If A[] is empty
          Output "the formula is unsatisfiable";
Else
Output the assignments in A[] as the satisfying assignments for the formula F.

Procedure getDistinctClauses(F[m])
Input: Formula with m clauses
Output: arrayofdistinctclauses and arrayofnondistinccaluses

distinctclauses[1] = clause[1];
n1:=0;
n2:=1;
distinct = true;
for i = 2 to m
          for j = 1 to distinctclauses.length – 1
                    if (distinctclause[j] intersection clause[i] != empty)
                              nondistinctclauses[n1++] = clause[i];
                              distinct = false;
                              break;
                    Endif
          Endfor
          If (distinct == true)
                    disticntclauses[n2++] = clause[i];
          Endif
Endfor

Procedure: removeSubsumedAssignments(tempassignments[], arraysubsumed[])
Input: list of assignments containing subsumed assignments and list of assignments subsuming the subsume assignments.
Output: list of assignments without subsumed assignments.

n:=0;
For i = 0 to tempassignments.length – 1
          For j = 0 to arraysubsumeb.length – 1
                    If (arraysubsumed[j] is not contained in tempassignents[i])
                              arrayassignments[n++] = tempassignment[i] ;
          Endfor
Endfor
Return  arrayassignments[];

Procedure: generateAssignment(lit, A[], temp[]);
Input: a literal in a clause and a list of assignments in A[].
Output: a list of assignments stored in temp[] produced by combining lit with A[].

For i = 1 to A.length
          If lit did not conflict with the assignment then
                    Combine the lit and the assignment;
                    Add the combination in temp[];
          elseif lit agrees with the assignment then
                    Add the assignment in temp[];
                    Add the assignment in arraysubsumed[];
          Endif
Endfor
Return temp[];
```

Figure 3 : The Proposed Algorithm

IV. Time Complexity of the Algorithm

The first three steps of the algorithm take polynomial time of number of clauses. Steps four and five are clearly the main contributors to the time complexity of the whole algorithm. These two steps rely on the number of assignments generated in each iteration of the for-loop. For step four, that number is determined by the clauses in CLS and for step five, that number is determined by the end of step four. Therefore, let us start with step four.

a) Finding number of assignments

Whenever a clause is considered in the for-loop, the number of assignments is multiplied by k (in the worst case). The first clause initializes A with k assignments (the literal in that clause). Then, the second clause will produce at most k^2 assignments, and the third clause may generate as k^3 assignments and so on. That means the number of the assignments is $\leq k^m$ where m is some number of clauses. In step four, clauses in CLS could either be:

1. $\frac{2n}{k}$ clauses (worst case).
2. or more than $\frac{2n}{k}$ clauses (as explained in Section 2).

Case 1 is the worst because if more than $\frac{2n}{k}$ clauses are needed, then we must have repeated literals. This can be shown easily as follows: If we have $\frac{2n}{k}+1$ clauses, then the number of literals is $(\frac{2n}{k}+1)k$ which gives us $2n+k$ literals. That means we have *k* repeated literals in these clauses.

Because of the existence of repeated literals in Case 2, Case 1 will produce the maximum number of assignments (see Theorem 3).

We now determine the number of possible assignments, $A(n)$, in the worst case. If the clauses in CLS have conflicting literals, $A(n) \leq k^m$.

In this case, a literal in one clause will not be combined with a literal $\bar{x}_1$ in another clause. The number of substitutions to be eliminated is shown by Theorem 2.

To count the exact number of assignments, the principle of inclusion-exclusion is used. The principle states that the number of elements that have property 1, property 2, property 3, …, or property n is found by the summation.

$$N(P_1,P_2,P_3,...,P_n)=\sum_{1\leq i\leq n}|A_i|-\sum_{1\leq i<j\leq n}|A_i\cap A_j|+\sum_{1\leq i<j<k\leq n}|A_i\cap A_j\cap A_k|-...+(-1)^{n+1}|A_1\cap A_2\cap A_3\cap...\cap A_n|$$

The proof of the principle can be found in (Rosen, 1999).

If P_i is the assignment where x_i and $\bar{x}_i$ appear for $i=1,2,3,...,\alpha$ where $\alpha \leq n$, then the exact number of assignments for case 1 is $k^{\frac{2n}{k}}-N(P_1,P_2,P_3,...,P_\alpha)$.

For any satisfiability instance, the previous quantity cannot be found. That is because unlike the example given earlier, the arrangement of variables or literals differs from one instance to another. However, there is an arrangement that will produce the highest number of variables.

b) The upper bound

At this point, we need to prove two theorems. One that states case 1 is the worst case and the other states the arrangement that will produce the highest number of assignments.

Theorem 3

In step 5 of the algorithm, generating assignments with the least number of clauses $(\frac{2n}{k})$ that include 2n literals is the worst case.

Proof

If more than $\frac{2n}{k}$ clauses are needed to include the $2n$ literals then we must have literals that are repeated. If we have one additional clause, then there must be k literals repeated and this will make the set of assignments to be excluded more than n. Having a repeated literal means that we have three clauses of this form: $x_1 \vee x_2 \vee x_3$ $x_1 \vee x_4 \vee x_5$ $\bar{x}_1 \vee x_6 \vee x_7$.

The two clauses that have the repeated literal x_1 will produce the unnecessary assignments. These assignments are generated when the repeated literal is combined with the $(k-1)$ literals of the other clause. This means that the assignments that include $\{x_1x_4, x_1x_5, x_1x_2, x_1x_3\}$ are unnecessary. The only useful assignment is x_1 produced from (x_1x_1). This indicates that $2(k-1)$ sets of assignments should be discarded. In addition to these assignments, the two repeated literals when combined with $\bar{x}_1$ will produce two sets of assignments of the form $(x_1\bar{x}_1)$ that are

also discarded from the total number of assignments when we count them using the inclusion exclusion principle. Therefore, a repeated literal will result to discard $2(k-1)+1$ additional sets excluded.

Writing the inclusion exclusion series with n sets plus $k(2(k-1)+1)$ sets is hard because there will be many possibilities for the intersection of sets. The approach to show that 2n/k is the worst case is to exclude the additional sets first from the total number of assignments and compare that with the worst case. The number of assignments of the additional sets can be counted by:

$$A=(2(k-1)+1)*C(k,1)*k^{\frac{2n}{k}-1}-(2(k-1)+1)^2*C(k,2)*k^{\frac{2n}{k}-3}$$
$$+(2(k-1)+1)^3*C(k,3)*k^{\frac{2n}{k}-5}-\ldots+(-1)^{k+1}(2(k-1)+1)^k*C(k,k)*k^{\frac{2n}{k}-2k+1}$$
$$=\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{\frac{2n}{k}-2i+1}$$

Excluding this from the total assignments

$$N=k^{\frac{2n}{k}+1}-\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{\frac{2n}{k}-2i+1}$$
$$N=k^{\frac{2n}{k}-2k+1}(k^{2k}-\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{2(k-i)})$$

Evaluating $(k^{2k}-\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{2(k-i)})$ for values of k gives quantity less than k^{2k-1} and result to a number of assignments less than $k^{\frac{2n}{k}}$ and excluding the n sets of the form (v -v) from N gives a value that is less than the one in the worst case.

$k^{\frac{2n}{k}}$ excld(n sets) $>k^{\frac{2n}{k}-2k+1}(k^{2k}-\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{2(k-i)})$ excld(n sets) because $k^{\frac{2n}{k}}>k^{\frac{2n}{k}-2k+1}(k^{2k}-\sum_{i=1}^{k}(-1)^{i+1}(2(k-1)+1)^iC(k,i)k^{2(k-i)})$.

This is for one additional clause. For i additional clauses the limit of the summation is to ik and this also will give the same result.

Theorem 3 tells us that step six will not generate assignments that are more than step five. This should make step 5 the dominant factor for time complexity.

Theorem 4

For the worst case, the upper bound is $(k(k-1))^{\frac{n}{k}}$

Proof

The inclusion-exclusion principle takes care of assignments that are counted more than once by considering the intersections between the n sets to be excluded as seen in the summation. Therefore, the least value of $N(P_1,P_2,P_3,\ldots,P_n)$ indicates the maximum possible number of assignments generated by the algorithm. This occurs when each set $x,\bar{x}$ is intersected with the maximum possible number of other sets. For example consider two clauses that has x_1 and $\bar{x}_1$ literals:

$$\begin{array}{lll} x_1 & x_2 & x_3 \\ \bar{x}_1 & x_4 & x_5 \end{array}$$

The assignments that include x_1 and $\bar{x}_1$ can never occur with assignments that include x_2 and $\bar{x}_2$, x_3 and $\bar{x}_3$, x_4 and $\bar{x}_4$, x_5 and $\bar{x}_5$ literals. Therefore, there is no intersection between $x_1\,\bar{x}_1$ assignment set and 4 sets of assignments. The least intersection $(k-1)$ happens if both clauses of x_5 and $\bar{x}_1$ have literals of the same variables. For the previous example the two clauses should look like this

$x_1 \vee x_2 \vee \bar{x}_3$ $\bar{x}_1 \vee \bar{x}_2 \vee x_3$ to make the quantity $N(P_1, P_2, P_3, ..., P_n)$ the least. If this happens with all variables, the following arrangement will produce the maximum number of assignments.

$$x_1 \vee x_2 \vee \bar{x}_3, \bar{x}_1 \vee \bar{x}_2 \vee x_3, \bar{x}_4 \vee \bar{x}_5 \vee x_6, x_4 \vee x_5 \vee \bar{x}_6, x_7 \vee \bar{x}_8 \vee x_9, \bar{x}_7 \vee x_8 \vee \bar{x}_9, \cdots$$

The number of assignments between clauses of conflicting literals is $k(k-1)$. Since we need $\frac{2n}{k}$ clauses to consider n variables and each 2 clauses have $k(k-1)$ assignments, then the number of assignments will be $k(k-1)^{\frac{n}{k}}$.

c) Related Work

Complete algorithms for SAT satisfiability problems include those algorithms that can state whether or not a SAT instance is satisfiable, giving a 'yes' answer together with a satisfying assignment or a 'no' answer as the case may be. The first complete algorithm is the Davis Putnam procedure (Davis & Putnam, 1960). This procedure is based on resolution rule to eliminate variables one by one till the formula is satisfied. When a variable is eliminated in each iteration, all resolvents are added to the set of the clauses. This algorithm requires polynomial space. It handles CNF formulas and it is one of the efficient SAT algorithms. (Davis, Logemann, & Loveland, 1962) Developed a divide-and-conquer algorithm that enhances on (Davis & Putnam, 1960). This improved algorithm is the main procedure for most state-of-the-art SAT solvers today. The search space of DPLL could grow as large as 2^n which is the worst case for any complete algorithm. Due to the possibility of consuming huge amount of time, researchers have been focusing on mechanisms to reduce that and came up with more reasonable time complexities. These improvements usually come in two aspects: the decision to branch to next literal and the backtracking mechanism if a solution is not found in the chosen branch. The achievements accomplished in improving SAT algorithm in these two aspects show that the complexity could be reduced significantly.

i. *Branching Decisions*

DPLL procedure chooses any literal for branching and goes down that region in the search space. The procedure will spend time searching for a solution and if it discovers that the branch is not successful, it backtracks to the other branch and continues searching. Choosing the next literal for branching more carefully will allow the algorithm to save time exploring a region where a satisfying assignment cannot be found at all and direct the searching to regions where a solution is likely to be found. In order to accomplish this, several heuristics have been proposed and the most effective ones can be found in (Bruni & A., 2003; Freeman, 1995; Hooker & Vinay, 1994; Jeroslow & Wang, 1990; Li & Anbulagan, 1997; Moskewicz, Madigan, Zhao, Zhang, & Malik, 2001; Pretolani, 1993).

ii. *Backtracking Mechanisms*

When the algorithm fails to find an answer or an empty clause (contradiction) appears down the path of the search tree, it backtracks to a certain point and continues searching in another part of the tree. The DP backtracks to the most recently untoggled (complemented) literal and tests its complement branch. As mentioned earlier this will cost a lot of time for DP to discover that this part of the search space does not have a solution and search for a solution elsewhere. For backtracking in the DP procedure, much work has not been done as compared to branching decision. This is due to the fact that backtracking is an essential step in any algorithm to prove its completeness. Nevertheless, there are a number of proposals to improve the backtracking in the DP procedure. (Lynce & Marques-Silva, Building State-of-The-Art SAT Solver, 2002) tested different backtracking strategies and the most effective ones can be found in (Lynce & Marques-Silva, 2002; Stallman & Sussman, 1977).

iii. *Upper Bounds*

The improvements made in backtracking and branching heuristics are of practical interests. However, the experimental analysis of these improvements indicates that satisfiability could be solved in time less than 2^n. A number of people gave lower bounds for this problem but most of them rely on a certain structure or property that exists in the formula. The following are some of the achievements made to find an upper bound that is better than the trivial one.

a. *Autarkness Principle*

The first attempt to achieve a non-trivial upper bound for SAT was done by (Monien & Speckenmeyer, 1985). They introduced the notion of autarks which are partial assignments of variables. If all clauses that include the variables in the assignment are satisfied, then that assignment is an autark. They proved that the time complexity of their algorithm is $O(2^{n \log \alpha_k})$.

b. *2-clause*

When dealing with 3-SAT problem, the clauses with 2 literals help in reducing the search space. Schiermeyer was the first to make use of the number of clauses with 2 literals after the resolution step is made (Schiermeyer, 1993). He said that for the next branch, a 2-clause is chosen such that it produces at least one new 2-clause in every branch that follows. With the help

of these reduced clauses, he proved an even lower bound for 3-SAT with time complexity $O(1.579^n)$. (Kullmann, 1999) showed that the algorithm of Schiermeyer can perform better through a new concept called blocked clauses. A clause C is blocked for a literal l if every clause C' containing l has also another literal that is complemented with C. By making use of these blocked clauses, Kullmann proved that the algorithm in (Schiermeyer, 1993) can have a time complexity of $O(1.504^n)$.

c. *Saisfiability Coding Lemma*

This lemma is based on isolated assignments which are satisfying assignments to the formula where a change of one value of any variable will make it dissatisfying. The lemma states that such assignments can be encoded in a message of length $(n-\frac{n}{k})$ and this is where the complexity comes from. (Marques-Silva & Sakallah, 1999) shows that through satisfiability coding lemma their algorithm finds an answer in $O(2^{n-\frac{n}{2k}})$.

d. *P-literal*

(Hirsch, Two New Upper Bounds for SAT, 1998) presented two algorithms that rely on P-literal notion. This notion says that if a literal occurs exactly 2 times in the clause set and at least 3 times in its negation form, then it is P-literal. He used these special literals to simplify the formula and came up with two algorithms with time complexity $O(2^{0.3089m})$ and $O(2^{0.10537L})$ respectively where m is the number of clauses and L is the length of the formula. An improvement was made to the second algorithm in (Hirsch, 2000) to become $O(2^{0.10299L})$.

e. *Covering Codes*

(Danstin, et al., 2002) proposed a deterministic algorithm that is based on covering codes. This algorithm can be seen as a derandomization of (Schoning, 1999) algorithm that uses random walk model. The search space is divided into group of assignments say balls of some radius r. Each group or ball represents some assignment *a* and all assignments that differ with it in r variables. The algorithm checks in each ball if there is a satisfying assignment and if there is none in any ball then the formula is unsatisfied. The authors of (Danstin, et al., 2002) showed that the time complexity of this $O(2-\frac{2}{k+1})$ for k-SAT. For 3-SAT, they managed to further improve the algorithm by identifying useless branching and reduce the search space to come up with running time $O(1.481^n)$.

V. Conclusion and Future Work

The proposed does not require the clauses or the formula to have any specific structure to achieve a competitive upper bound which is a significant advantage over the existing algorithms in the literature where they derive their time complexity based on a property that must exist in the formula. The algorithm gives a new insight towards solving SAT. Most of the other algorithms are based on the classical rule of splitting the search space into regions and search for a solution in each one. The new perspective of the algorithm has the potential to design further effective SAT algorithms that outperforms the existing ones in theory and practice.

The implementation of the proposed algorithm will be considered in future work. The algorithm proposed here can also be improved. The time complexity of the proposed algorithm is based on pre-processing of clauses in the formula. This arrangement is so unlikely to exist in all clauses considered. That means that there exists a tighter upper bound for the algorithm but to achieve that the order in which clauses are considered should be more intelligent. To show that such an upper bound exists, many cases have to be covered and counted. Parallelisation of the proposed algorithm is also a potential future work.

References Références Referencias

1. Barbour, A.E. (1992). Solutions to The Minimization Problem of Fault-Tolerant Logic Circuits. *IEEE Transactions on Computers, 41(4),* 429-443.
2. Bruni, R., & A., S. (2003). A Complete Adaptive Algorithm for Propositional Satisfiability. *Discrete Applied Mathematics, 127*.
3. Cook, S. (1971). The Complexity of Theorem Proving Procedures. *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing,* (pp. 151-158).
4. Crawford, J. M., & Baker, A.B. (1994). Experimental Results on the Application of Satisfiability Algorithms to Scheduling Problems. *AAAI-94.*
5. Danstin, E., Goerdt, A., Hirsch, E. A., Kannan, R., Kleinberg, J., Papadimitriou, C., et al. (2002). A Deterministic (2–2 / k+1)n Algorithm fork-SAT based on Local Search. *Theoretical Computer Science,* 69-83.
6. Davis, M., & Putnam, H. (1960). A Computing Procedure for Quantification Theory. *Journal of Association for Computing Machinery,* 201-215.
7. Davis, M., & Putnam, H. (1960). A Computing Procedure for Quantification Theory. *Journal of Association for Computing Machinery,* 201-215.
8. Davis, M., Logemann, G., & Loveland, D. (1962). A Machine Program for Theorem Proving. *Communications of the ACM, 5,* (pp. 394-397).

9. Devadas, S. (1989). Optimal Layout via Boolean Satisfiability. *Proceedings of ICCAD 89,* (pp. 294-297).
10. Freeman, J. W. (1995). *Improvements to Propositional Satisfiability Search Algorithms. Ph. D. Dissertation.* Department of Computer and Information Science, University of Pennsylvania.
11. Hirsch, E. (2000). New Worst Case Upper Bounds for SAT. *Journal of Automated Reasoning, 24,* pp. 397 – 420.
12. Hirsch, E. (1998). Two New Upper Bounds for SAT. *Proceedings of 9th Annual ACM Siam Symposium on Discrete Algorithms* (pp. 521 – 530).
13. Hooker, J. N., & Vinay, V. (1994). Branching Rules for Satisfiability. *GSIA Working Paper 1994-09.* Pennsylvania: Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh .
14. Jeroslow, R. G., & Wang, J. (1990). Solving Propositional Satisfiability Problems. *Annals of Mathematics & Artificial Intelligence, 1,* 167-187.
15. Kauts, H., & Selman, B. (1992). Planning as Satisfiability. *Proceedings of the 10th European Conference on Artificial Intelligence (ECAI 92).*
16. Kullmann, O. (1999). New Methods for 3-SAT Decision and Worst Case Analysis. *Theoretical Computer Science,* 1-72.
17. Larrabee, T. (1992). Test Pattern Generation Using Boolean Satisfiability. *IEEE Transactions Computer Aided Design, II(1),* 4-15.
18. Li, C. M., & Anbulagan. (1997). Heuristics Based on Unit Propagation for Satisfiability Problems. *Proceedings of 15th International Joint Conference on Artificial Intelligence, 1,* pp. 366-371. Nagoya, Japan.
19. Lynce, I., & Marques-Silva, J. (2002). Building State-of-The-Art SAT Solver. *Proceedings of the European Conference on Artificial Intelligence (ECAI) 105.*
20. Lynce, I., & Marques-Silva, J. P. (2002). *The Effect of Nogood Recording in MAC-CBJ SAT Algorithms.* Technical Report RT/4/2002, INESC.
21. Marques-Silva, J. P., & Sakallah, K. A. (1999). IEEE Transactions on Computers. *GRASP-A Search Algorithm for Propositional Satisfiability,* 506 -521.
22. Monien, B., & Speckenmeyer, E. (1985). Solving Satisfiability in less than 2n steps. *Discrete Applied Mathematics,* 287-295.
23. Moskewicz, M., Madigan, C., Zhao, Y., Zhang, L., & Malik, S. (2001). Engineering an Efficient SAT Solver. *Proceedings of the Design Automation Conference.*
24. Pretolani, D. (1993). Efficient and Stability of Hypergraph SAT Algorithms. *Proceedings of DIMACS Challenge II Workshop.*
25. Rosen, K. H. (1999). *Discrete Mathematics and Its Applications* (4th Edition ed.). McGraw Hill.
26. Schiermeyer, I. (1993). Solving 3-Satisfiability in less than 1.579n. *In Selected papers from Computer Science Logic 12, LNCS 702,* (pp. 379-394).
27. Schoning, U. (1999). A Probabilistic Algorithm for k-SAT and Constraint Satisfaction Problems. *Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS)* (pp. 410-414). IEEE.
28. Stallman, R. M., & Sussman, G. J. (1977). Artificial Intelligence 9. *Forward Reasoning & Dependency-directed Backtracking in A System for Computer-aided Circuit Analysis,* (pp. 135-196). Artificial Intelligence 9.