Improving NLSAT for Nonlinear Real Arithmetic Theory with Clause-Level Information

Anonymous Author(s)*

ABSTRACT

Model-constructing satisfiability calculus (MCSAT) framework has been applied to SMT problems on different arithmetic theories. NLSAT, an implementation using cylindrical algebraic decomposition for explanation, is especially competitive among nonlinear real arithmetic constraints. However, current Conflict-Driven Clause Learning (CDCL)-style algorithms only consider literal information for decision, and thus ignore clause-level influence on arithmetic variables. As a consequence, NLSAT encounters unnecessary conflicts caused by improper literal decisions.

In this work, we analyse the literal decision caused conflicts, and introduce clause-level information with a direct effect on arithmetic variables. Two main algorithm improvements are presented: clause-level feasible set based look-ahead mechanism and arithmetic propagation based branching heuristic. We implement our solver named clauseSMT on our dynamic variable ordering framework. Experiments indicate that clauseSMT is competitive on nonlinear real arithmetic theory against existing SMT solvers (CVC5, Z3, YICES2), and outperforms all these solvers on satisfiable instances of SMT(QF_NRA) in SMT-LIB. The effectiveness of our proposed methods are also studied.

CCS CONCEPTS

• **Theory of computation** → **Logic and verification**; **Automated reasoning**; *Constraint and logic programming*; *Automated reasoning*; • **Mathematics of computing** → **Solvers**.

KEYWORDS

NLSAT, Nonlinar Real Arithmetic, SMT, Clause Level

1 INTRODUCTION

Satisfiability Modulo Theories (SMT) is a kind of problem which aims to detect the satisfiability of formulas in first order logic. SMT problems are usually involved in theories like linear and nonlinear arithmetic, uninterpreted functions, strings and arrays [8]. As a fundamental problem in software engineering, formal method and programming languages, it is widely used in various applications, such as symbolic execution [10, 25], program verification [9, 53], program synthesis [50], automata learning [48, 52] and neural network verification [4, 31, 45, 46].

Nonlinear real arithmetic (NRA) is one kind of arithmetic theories. It contains atoms which are represented as inequalities of polynomials, thus sometimes it is called theories about polynomial constraints. Variables can take boolean values or real numbers according to their types. Instances in SMT(NRA) are usually generated from academical and industrial applications. It is commonly used in cyber physical systems [5, 15, 47], ranking function generation [26, 35] and nonlinear hybrid automata analysis [18]. Instances from these applications are collected in the SMT-LIB benchmarks [7]. All these applications gain improvement thanks to the high performance of SMT solvers over nonlinear arithmetic.

Decision procedures about nonlinear arithmetic solving are usually based on cylindrical algebraic decomposition (CAD) [13], a useful real quantifier elimination tool. CAD is used to generate the current unsat cell during the search procedure, and is employed in modern algorithms. Among them, NLSAT [30] is the mainstream algorithm which uses CAD for lemma generation. Its basic idea is to directly assign the value to arithmetic variables, rather than on literal levels, as CDCL(T) does.

Although NLSAT designs a novel view of directly assigning arithmetic variables, it still maintains literal decisions when processing arithmetic clauses with several unevaluated literals. Decided literals are then used for conflict analysis in such a CDCL-style framework. However, improper literal decisions sometimes cause conflicts and thus decrease the speed of the whole search procedure. Thus, a heuristic for literal decisions is required.

We present three central problems and give our solutions to them during the algorithm improvement.

- What causes conflicts during NLSAT algorithm? Can we avoid some of them?
- Is there any possible to directly assign the arithmetic variable, regardless of literal-level decision information, in such a CDCL-style framework?
- Can we do propagation on arithmetic variables, just like unit propagation in SAT solving? Is the new propagation method helpful for new assignment and conflict detection?

To answer the questions above, this paper for the first time proposes a new algorithm considering clause-level information. First, we analyze the conflict problems caused in NLSAT, and divide them into two categories. As described in [36], each clause narrows the feasible set[1] of an arithmetic variable. This technique has been used to enlarge the operation choices in local search algorithm. However, it has not been considered when designing complete methods like NLSAT. Arithmetic variables are sometimes narrowed to an empty search space, and thus conflict happens. We describe this kind of problem in the view of interval arithmetic, and gives the solution thanks to the computation of feasible intervals. The clause-level feasible set idea extends the spirit of NLSAT that directly makes progress on arithmetic variables. Second, we introduce the incremental computation of clause-level feasible set, followed by the definition of clause-level propagation. In SAT solvers, unit propagation is an effective tool to deduce the assignment and detect conflict clauses. Similarly, we adopt the clause-level propagation to fix a possible witness for the arithmetic variable, or detect the empty feasible set cases as quickly as possible. Finally, we also present the structure and implementation details of our solver clauseSMT. Although dynamic variable ordering has been discussed in [43], SMT-RAT [19] solves fewer instances due to the lack of effective

[1] Also called satisfying domain in [36].

data structures. We present our techniques and data structures borrowed from SAT solving. We implement the above ideas in the NLSAT module of Z3 solver [20]. Some mathematical methods like root isolation, polynomial operations and algebraic number representation are relied on the existing libraries in Z3. We perform the experiments on the SMT-LIB benchmark, and study the effect of proposed techniques, including look-ahead mechanism and clause-level propagation. The experiment results show that our algorithm solved the most satisfiable instances and is overall very competitive against other SMT solvers.

The structure of the paper is organized as follows: we first give the definition of SMT problems on nonlinear real arithmetic, and also the traditional complete method NLSAT in Section 2. Analysis of conflicts in NLSAT algorithm is studied in Section 3. A feasible set based look-ahead mechanism is presented in Section 4. Followed by the idea of clause-level information, we present the algorithm for clause-level propagation and a new branching heuristic in Section 5. In Section 6, we discuss the implementation details. We compare our solver with other SMT solvers and perform ablation study in Section 7. Related works about nonlinear real arithmetic solving are supplemented in detail in Section 8. Finally, we conclude our work and propose potential research directions in Section 9.

2 PRELIMINARIES

In this section, we introduce the basic definition of SMT problems over nonlinear real arithmetic theory, followed by NLSAT algorithms. Besides, we also introduce the computation of clause-level feasible set.

2.1 Syntax of SMT(QF_NRA)

The syntax of SMT constraints on nonlinear real arithmetic is defined as follows:

$$\begin{aligned} \text{variables:} &\quad v := x \mid b \\ \text{polynomials:} &\quad p := x \mid c \mid p + p \mid p \cdot p \\ \text{atoms:} &\quad a := b \mid p \le 0 \mid p \ge 0 \mid p = 0 \\ \text{literals:} &\quad l := a \mid \neg a \\ \text{formulas:} &\quad c := l \mid l \vee l \mid l \wedge l \end{aligned}$$

Here we denote x for an arithmetic variable which takes real numbers, b for a boolean variable which takes true or false. In nonlinear real arithmetic, an atom is either a boolean atom, defined by a boolean variable, or an arithmetic atom, defined by a strict or non-strict inequality of a polynomial. A literal is considered to be an atom or its negation. A clause is defined by the disjunction combination of literals. In our method, input instances are all transformed into conjunctive normal form (CNF), where the whole constraint is represented by the set of clauses. We denote $\mathbb{B}$ as a set of boolean variables, $\mathbb{V}$ as a set of arithmetic variables, SMT(NRA) as formulas on the theory of nonlinear real arithmetic.

For the semantics, we define an assignment α as a mapping from variables to values. Specifically, a boolean assignment maps boolean variables to boolean values, written as $\alpha(bool) : b \to \{\top, \bot\}$. An arithmetic assignment maps arithmetic variables to real values, written as $\alpha(real) : v \to \mathbb{R}$. In complete methods like MCSAT, a complete assignment maps all boolean and arithmetic variables to values, while an incomplete assignment only maps part of them. Under a given assignment, an atom can be evaluated to be true if the assignment satisfies it, false if the assignment breaks it, undefined if the atom contains a variable which is not mapped by the assignment. A complete assignment satisfying all clauses is also called a solution to the given instances, and thus proves it is satisfiable. The SMT problem on nonlinear real arithmetic is to look for a solution for the given instance, or prove that it can not be satisfied by any assignment.

2.2 Feasible Set

For nonlinear real arithmetic constraints, a key technique used for determining the possible values of arithmetic variables is called root isolation. Given an arithmetic atom $p\ \{\le, \ge, =, >, <\}\ 0$, if there is only one variable v not assigned by the assignment, we can calculate the set of its possible values satisfying the atom, called feasible set. In high order polynomial constraints, the computation of feasible set is usually based on root isolation, which uses mathematical methods to solve the roots of the polynomial. Intervals on the real space separated by the roots[2] maintain the evaluation value for the atom, and thus their combination represents the overall (in)feasible set. For literals in a negation form, the feasible set of the literal is actually the complement of the corresponding feasible set of the atom. By forcing the arithmetic variable any value in the feasible set, the atom will be satisfied[3].

Besides the literal level, feasible set can be extended to the clause level, which represents the set of values to make the clause satisfied. We also consider the circumstances when there is only one arithmetic variable being unassigned in the clause. We call that kind of clause a univariate clause. The formal definition of feasible set is shown as follows:

Definition 2.1. Given a literal l, an arithmetic variable x, an assignment assigning all variables appearing in the literal l but x, the *feasible set* (resp. *infeasible set*) is the combination of intervals, in which the literal l is satisfied (resp. unsatisfied) when x is mapped.

Similarly, given a clause c, an arithmetic variable x, an assignment assigning all variables appearing in the clause c but x, the *feasible set* (resp. *infeasible set*) means that the clause c is satisfied (resp. unsatisfied) when x is mapped to any value in it. The computation of *feasible set* (resp. *infeasible set*) is done by just taking the union (resp. intersection) of computed feasible set (resp. infeasible set) in respect to all literals in that clause. An example of a feasible set is shown in Example 2.2.

Example 2.2. Given an assignment $\alpha := \{b \to \top, x \to 4\}$, the feasible set of the clause below

$$b \vee y + x > 0 \vee y^2 > 2$$

should be $\{(-4, -\sqrt{2}) \cup (\sqrt{2}, \infty)\}$

2.3 Original NLSAT Algorithm

NLSAT is the main part algorithm over nonlinear real arithmetic theory inside Z3 solver [20]. NLSAT have the ability to handle both

[2]Intervals are also called cells in the perspective of cylindrical algebraic decomposition.
[3]Sometimes the feasible set can be an empty set or a full set, which means the atom is always satisfied or unsatisfied by choosing any value for the arithmetic variable respectively.

Algorithm 1: Process Clauses in NLSAT

```
Input  : A set of clauses F
Output : Conflict clause conf_cls
for clause c in clauses F do
    for literal l in clause c do
        feasible_set ← compute feasible set of literal;
        literal_value ← real propogate literal;
        if literal_value = ⊤ then
            break;
        if literal_value = ⊥ then
            continue;
    if only one literal undefined then
        unit propagate;
    else if two or more literals undefined then
        decide the first literal;
    else
        return c;
return No Conflict;
```

Algorithm 2: Original NLSAT

```
Input  : A formula F
Output : SAT or UNSAT
while true do
    variable v ← branching heuristic;
    conf_cls ← process clause set univariate to v;
    if conf_cls is empty then
        // Consistent
        if next variable is boolean then
            Boolean decide;
        else if next variable is arithmetic then
            Semantics decide;
        else
            // No unassigned variables
            return SAT;
    else
        // Conflict
        new_lemma ← conflict analysis;
        if new_lemma is empty then
            return UNSAT;
        else
            backtrack;
```

boolean and arithmetic variables directly. In other words, it incorporates the theory reasoning into CDCL framework, by extending unit propagation and boolean decisions into real-propagation (R-propagation) and semantic decisions.

To select the right value for an arithmetic variable, NLSAT updates the feasible set of them in the search process. Assume the current feasible set is $curr\_set$, literal $lit's\ feasible\_set$ is lit_set, real-propagation takes effect when one of the circumstances below happens.

- **lit_set is empty**: the literal is propagated false because it can't be satisfied by any value.
- **lit_set is full**: the literal is propagated true because it can be satisfied with whatever value we choose.
- **curr_set is a subset of lit_set**: the literal is propagated true, since the current feasible set has already satisfied it.
- **curr_set has no intersection with lit_set**: the literal is propagated false, since choosing values from current feasible set must make the literal unsatisfied.

When processing a clause with boolean and arithmetic variables, NLSAT first uses propagation methods and evaluation to detect evaluated literals. If one of them is true, the clause is skipped; otherwise, NLSAT decides the first literal and updates the feasible set. Algorithm 1 shows the processing process section in NLSAT. For the conflict analysis, NLSAT uses cylindrical algebraic decomposition (CAD) as a tool for explanation. Model-based projection generates the conflict cell and learns a lemma to avoid entering it in the future. Algorithm 2 shows a complete NLSAT method.

3 CONFLICTS DURING NLSAT ALGORITHM

In this section, we analyze the reasons that cause conflicts in NLSAT algorithms. Generally, they can be divided into two categories, semantics decision caused conflicts and literal decision caused conflicts.

3.1 Conflicts Caused by Semantics Decision

In SAT solving, the conflicts are always caused by literal decisions (or boolean variable decisions). For a satisfiable instance, SAT solvers can avoid any conflicts when using a perfect phase selection (i.e. assign variables true or false) method. For unsatisfiable instances, conflicts are encountered whatever phase the solver selects. In both cases, when CDCL detects a conflict by wrong decision values, conflict analysis is invoked to generate a new lemma which forces the previous assignment to change.

Similarly, in NLSAT algorithm for SMT solving, conflicts are caused by wrong semantics decision, i.e. choosing a value from a given interval. As discussed in [39], the search space of nonlinear arithmetic is composed by sign-invariant cells. However, in systematic solving like NLSAT, we cannot forecast the cell we are currently searching, and thus conflicts may happen. For an unsatisfiable instance, each cell in the search space is inconsistent with all polynomial constraints, and thus conflicts can not be avoided. We give the demo in Example 3.1.

Example 3.1. Consider the formula is $\{y^2 + x + 1 \leq 0\}$, the variable order is $\{x, y\}$. As depicted in Figure 2, when we decide $\{x \to 2\}$, the satisfying area (shaded area) has no intersection with the dashed line $x = 2$, and thus a conflict happens. This kind of conflicts is caused by wrong semantics decision value for previous variable x, and can be avoided if we choose a right value, for example $\{x \to -2\}$.

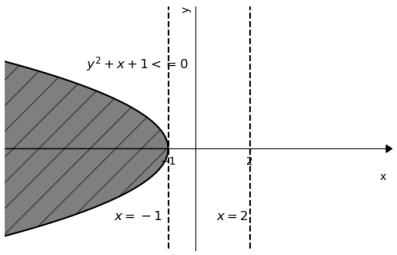

Figure 1: Demo of semantics decision caused conflict.

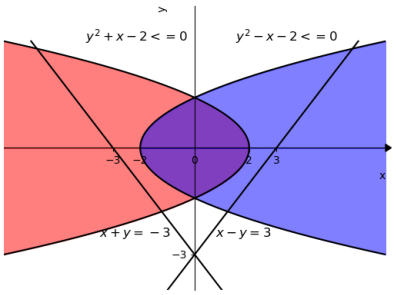

Figure 2: Demo of literal decision caused conflict.

3.2 Conflicts Caused by Literal Decision

A key technique used in NLSAT is processing clauses univariate to the current arithmetic variable. In CDCL-style systematic search, literals are required to be assigned by unit propagation for unit clauses (only one literal unassigned), or decisions for clauses with multiple unassigned literals. However, there has been less attention paid to the literal decision mechanism of NLSAT. Improper literal decisions may cause extra conflicts. We explain this circumstance in Example 3.2.

Example 3.2. Suppose we have three clauses shown below:

$$c_1 : y^2 + x - 2 \leq 0 \vee y^2 - x - 2 \leq 0$$
$$c_2 : x + y = -3 \quad c_3 : x - y = 3$$

As depicted in Figure 2, the purple area satisfies both polynomials in the clause c_1, while the red area and blue one only satisfy $y^2+x-2 \leq 0$ and $y^2 - x - 2 \leq 0$ respectively. Two straight lines represent two equality constraints, respectively. Suppose the formula for the SMT problem is $\{c_1, c_2\}$, the intersection is located only in the red area. If the formula is $\{c_1, c_3\}$, the intersection is then located only in the blue area. However, when NLSAT algorithm processes a clause with multiple unassigned literals like c_1, it will decide one literal and branch the search space into either red+purple area, or blue+purple area. In this case, there is a 50 percent possibility to miss the straight line and make conflict happen. However, it is noticed that the feasible set of clause c_1 (i.e. the red+blue+purple area) has a nonempty intersection with both straight lines, which means the formula $\{c_1, c_2\}$ and $\{c_1, c_3\}$ are both satisfiable.

Here comes a new problem for this circumstance. Is it possible to avoid conflicts by a better literal decision heuristic? If so, how can we do this? The key idea is to view the literal decision problem as a satisfiability about intervals. We give another Example 3.3 to help explain.

Example 3.3. Suppose the current assignment is $\alpha ::= \{x \rightarrow 0\}$.Clauses univariate to y are shown as follows:

$$c_1 : \quad (y+2)(y+4) \leq x \vee (y-2)(y-4) \leq x$$
$$c_2 : \quad (y+5)(y+6) \leq x \vee (y-1)(y-5) \leq x$$

By calculating the feasible set of each literal, the interval view of the clauses is shown below.

$$c_1 : [-4, -2] \vee [2, 4]$$
$$c_2 : [-6, -5] \vee [1, 5]$$

Then the problem can be demonstrated like this: is there a value that is included in at least one interval of all clauses?

4 FEASIBLE SET BASED LOOK-AHEAD MECHANISM

In this section, we enrol the clause-level feasible set into NLSAT algorithm and design a look-ahead mechanism.

4.1 Look-Ahead Before Processing Clauses

We first extend the feasible set definition to a clause set.

Definition 4.1. Given a clause set cs, an arithmetic variable x, an assignment assigning all variables appearing in any clause in cs but x, the *feasible set* (resp. *infeasible set*) means that each clause c is satisfied (resp. unsatisfied) when x is mapped to any value in it. Specifically, the feasible set can be computed by taking the intersection of each clause's feasible set.

$$feasible\_set(cs) = \bigcap_{c \in cs} feasible\_set(c)$$

By using the feasible set of a clause set, it's much easier to directly conclude the search space of an arithmetic variable. Concretely, when the feasible set is not empty, by assigning any value in the set to the variable (semantics decision), the search will make progress and jump to the next stage. On the contrary, when the set is empty, there is no way to avoid inconsistency by choosing a value for the variable. Now we answer the question proposed in Section 3 and give the formal definition below.

Definition 4.2. When the feasible set of the clause set is not empty, theoretically we can avoid conflicts by deciding proper literals, we call this case a *path case*. In contrast, when the feasible set is empty, the conflict can not be avoided by literal decisions (is caused by semantics decisions), we call this case a *block case*. Example 4.3 and 4.4 show a path case and a block case, respectively.

Example 4.3.

$$\left.\begin{array}{l} c_1 : [-4,-2] \vee \boxed{[2,4]} \rightarrow \{[-4,2] \cup [2,4]\} \\ c_2 : [-6,-5] \vee \boxed{[1,5]} \rightarrow \{[-6,-5] \cup [1,5]\} \end{array}\right\} \bigwedge \rightarrow \{[2,4]\}$$

This example shows a path case, since the clauses can be satisfied by deciding literals in the green boxes.

Algorithm 3: Deciding Literals using pre-appointed value

Input : A set of clauses F, pre-appointed value val selected from feasible set
Output : Decided literals $lits$

```
for clause c in clauses F do
    for literal l in clause c do
        feasible_set ← compute feasible set of literal;
        literal_value ← real propogate literal;
        if literal_value = ⊤ then
            break;
        if literal_value = ⊥ then
            continue;
        // decide satisfiable literals under val
        if feasible_set contains val then
            path_literal ← l;
    if only one literal undefined then
        unit propagate;
    else
        lits ← lits ∪ {path_literal};
return lits;
```

Example 4.4.

$$\left.\begin{array}{l} c_1 : \boxed{[-4,-2]} \vee \boxed{[2,4]} \rightarrow \{[-4,2] \cup [2,4]\} \\ c_2 : \boxed{[-6,-5]} \vee \boxed{[5,6]} \rightarrow \{[-6,5] \cup [5,6]\} \end{array}\right\} \bigwedge \rightarrow \varnothing$$

This example shows a block case, the clauses can not be satisfied whatever literals we decide.

The feasible set computation gives us a view of the current consistent search space. However, in such a CDCL-style algorithm, literals must be assigned for future conflict analysis. Here comes another question? How do we decide literals if we already know it's a path case (i.e. find green boxes in Example 4.3)? In our work, we use a look-ahead mechanism that allows us to select a pre-appointed value from the feasible set. Then we utilize the pre-appointed value to look for a consistent decision path. We give the detail in Algorithm 3.

The updated algorithm adds a condition that the feasible set of current literal should include the pre-appointed value. By doing so, feasible sets of all decided literals during the processing procedure will intersect with each other into a clause-set-level feasible set, which allows the arithmetic variable to be assigned by that appointed value. For the block case, the process algorithm is the same with NLSAT, which will later enter resolve procedure and force previous arithmetic assignments to change.

4.2 Look-Ahead After Conflict Analysis

Besides processing clauses with multiple literals, our decision-making algorithm still shows effectiveness for cylindrical algebraic decomposition (CAD) based explanation. CAD uses projection on conflict polynomials to learn a new lemma eliminating a sign-invariant cell. The generated lemma contains an extended

Algorithm 4: Process clauses after a new lemma

Input : A new lemma $lemma$, arithmetic variable v

```
lemma_feasible_set ← compute feasible set of clause(lemma);
feasible_set[v] = feasible_set[v] ∩ lemma_feasible_set;
if feasible_set[v] is empty then
    // block case
    // same with nlsat, call Algorithm 1
    original process clauses;
else
    // path case
    val ← value_selection(feasible_set[v]);
    // call Algorithm 3
    deciding literals using pre-appointed value (val);
```

polynomial constraint called root atom as shown below:

$$y \sim root_i(p(x_1, ..., x_n))$$

where y is the last assigned variable, $\sim \in \{=, \neq, \leq, \geq, <, >\}$, p is a polynomial generated using model-based projection, containing previously assigned variables $x_1, ..., x_n$. Specifically, when the last assigned variable y is located between two polynomial constraints, several root atoms are generated in the new learned lemma, and thus make the literal decision important.

Example 4.5.

$$\begin{array}{ll} c_1 : \boxed{[-7,-2]} \vee [2,8] & c_1 : [-7,-2] \vee \boxed{[2,8]} \\ c_2 : [-11,-10] \vee \boxed{[-6,5]} & c_2 : [-11,-10] \vee \boxed{[-6,5]} \\ learned : \boxed{[3,4]} \vee \boxed{[7,8]} & learned : \boxed{[3,4]} \vee [7,8] \end{array}$$

Before learning a new lemma, left shows a possible path in green boxes. However, the updated feasible set is not consistent with any literal in the lemma, which causes a conflict. Right shows a new possible case after the incremental lemma.

Look-ahead algorithm after conflict analysis is similar to the main search part, as shown in Algorithm 4. The main difference is the incremental computation of decision cases by considering only the feasible set of the new lemma. For the previously processed clauses, we use a vector of intervals to cache their feasible sets. By union the feasible set of the new lemma, we can quickly return the decision case of the arithmetic variable. For the path case, we must reprocess the clauses and try to find a new decision path because the new generated lemma adds a new constraint on the arithmetic variable, and sometimes forcing the current decision path blocked, as shown in Example 4.5.

5 CLAUSE-LEVEL PROPAGATION

Following the idea of using clause-level feasible set, this section introduces a new kind of propagation called clause-level propagation. The idea is inspired from unit propagation (or literal propagation) in SAT solving. For a boolean satisfaction problem, boolean variables can be unit propagated to assign a boolean value. By propagating boolean variables, sat solvers can assign an unassigned variable,

Algorithm 5: Clause-Level Propagation

Input: Clause set F

```
for clause cls do
    if cls is univariate to an arithmetic variable v' then
        cls_feasible_set ← compute_feasible_set of clause (cls);
        feasible_set[v'] = feasible_set[v'] ∩ cls_feasible_set;
        if feasible_set[v'] is empty then
            blocked_vars = blocked_vars ∪ {v'};
        if feasible_set[v'] is single value then
            fixed_vars = fixed_vars ∪ {v'};
```

or detect conflicts as soon as possible. However, currently most complete algorithms don't consider doing arithmetic propagation for quick assignment or conflict detection. Here we give our formal definition of clause-level propagation in Definition 5.1.

Definition 5.1. Given a clause c, an arithmetic variable x, an assignment α assigning all variables appearing in the clause but x. A clause-level propagation on x is to compute the feasible set of current clause c, which narrows the feasible set of variable x.

The difference between arithmetic and boolean problems is about the states of search spaces. Unit propagation always propagates the boolean variables to be true or false. In other words, unit propagation always prune the search space of the variable a half, and thus guides the search to proceed (cause there are only two possible values for boolean variables). However, in arithmetic view, clauses sometimes only prune the search space by part of the real space, and thus only make the search space contract. In this section, we introduce the algorithm of clause-level propagation, and then use the propagation information to guide our search by changing the next branching variable.

5.1 Clause-Level Propagation Method

In SAT solving, unit propagation happens after assignment. In our algorithm, we incrementally calculate the feasible set of a clause, when it's univariate to an arithmetic variable[4]. The new generated univariate clause will add an additional constraint to the arithmetic variable, and thus prune the search space by taking the intersection. Unlike the boolean search space, we divide the propagation cases into three kinds, as shown in Example 5.2:

- **Block case:** The clause-level feasible set is empty.
- **Fixed case:** The clause-level feasible set contains only one real number, for example $[2, 2]$.
- **Other case:** The clause-level feasible set is narrowed, but not empty or only include one real value.

This information is later used for better branching heuristic. We show the details in Algorithm 5.

[4]This does not only happen after the assignment of an arithmetic variable, but also after a boolean variable is assigned. Whatever, when the clause is arithmetically univariate, the feasible set is updated.

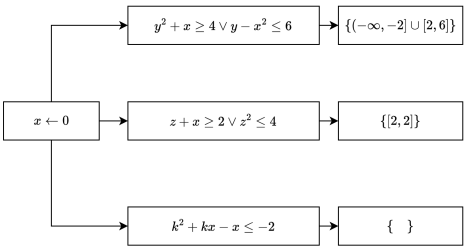

Figure 3: Demo of clause-level propagation for y (normal case), z (fixed case) and k (block case).

Algorithm 6: Branching Heuristic of Variables

Output: A variable v

```
// branch blocked variables first (conflict)
if blocked_vars ≠ ∅ then
    v ← select_from(blocked_vars);
// branch fixed variables next (propagate value)
else if fixed_vars ≠ ∅ then
    v ← select_from(fixed_vars);
else
    v ← vsids_select;
return v
```

Example 5.2. An example is shown as Figure 3. When a variable x is assigned to 0, three clauses become univariate to other variables $\{x, y, k\}$. These three clauses add three new constraints on arithmetic variables, calculated as $\{(-\infty, -2] \cup [2, 6]\}, \{[2, 2]\}, \emptyset$. These three feasible sets just indicate that variables are feasible, fixed or blocked.

5.2 Propagation-Based Branching Heuristic

After unit propagation in SAT solving, an unassigned variable is propagated to a value, or a conflict clause is detected effectively. When we get the feasible set for arithmetic variables, propagation and conflict cases still exist, which respectively match the fixed and blocked case talked above. However, current clause-level conflict cannot return a conflict clause directly in NLSAT, which is actually the responsibility of processing clauses algorithm. Thus, we use the recorded fixed and blocked information, and force the branching heuristic to aim at solving those variables as soon as possible.

For the normal case, we use Variable State Independent Decaying Sum (VSIDS) [40] as the main branching heuristic, as suggested in [29] and [43]. We give the branching heuristic in Algorithm 6. Details about our implementation of dynamic variable ordering is talked in Section 6.

6 IMPLEMENTATION

All the above algorithms are incorporated into our new solver called clauseSMT[5]. We describe the details of the solver, which include a detailed version of conflict analysis, and also implementations of

[5]We present our experimental results and binary file for Linux system on an anonymous repository https://anonymous.4open.science/r/ASE2024_clauseSMT/README.md

Algorithm 7: ClauseSMT

```
Input  : A formula F
Output : SAT or UNSAT
while true do
    clause level propagation (F);
    // call algorithm 5
    variable v ← branching heuristic;
    // call algorithm 6
    if v's feasible set is empty then
        new_lemma ← Resolve (Conflict Analysis);
        if new_lemma is empty then
            return UNSAT
        else
            Process Clauses after a new lemma
             (new_lemma);
            // call algorithm 4
    else
        val ← select from feasible set;
        Process Clauses using pre-appointed value val;
        // call algorithm 3
        assign v ← val;
        if all variables are assigned then
            return SAT;
```

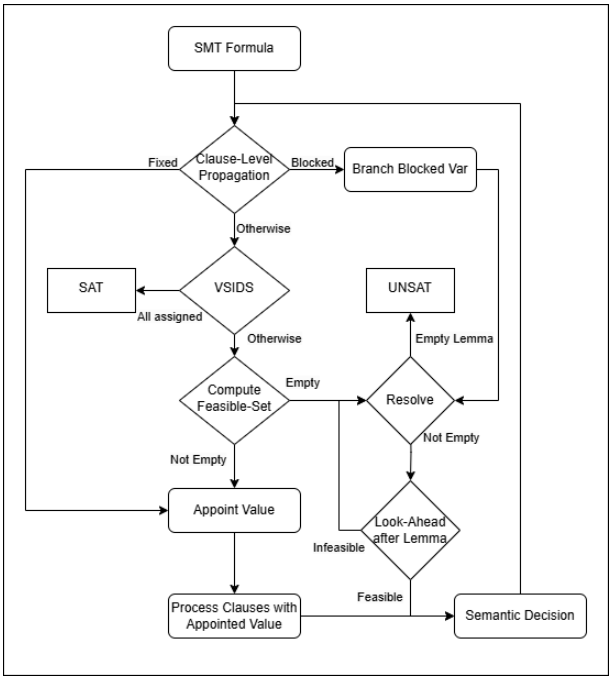

Figure 4: Overall Structure of clauseSMT.

dynamic variable ordering framework. The structure of clauseSMT is depicted in Figure 4. Algorithm 7 shows the integrated heuristics implemented in ClauseSMT.

6.1 Resolve

Although look-ahead based processing algorithm is given in previous section, cases are more complex for the implementation. As discussed, conflicts in NLSAT can be divided into two kinds. For the literal decisions case [6], we maintain the resolve algorithm same with original algorithm. The search engine should backtrack to the maximum decision level of the new lemma, and try to unit propagate the negation of previously decided literals. In other words, this case happens just in pure literal levels. For the second case, we detect literal decision cases with consideration of the incremental clause (learned lemma). Note that since we need to recalculate the decision path, literals that have been decided at that stage need to be reset.

To better management the backtrack part, we introduce several new kinds of trail normally used in NLSAT algorithm:

- **path_finder**: whenever we check the feasible set of current stage is non-empty, a path exists and this trail is saved.
- **block_finder**: when we find the current stage is blocked, we save this trail.
- **clause_feasible_updated**: whenever we update a clause feasible set of an arithmetic variable, the trail is saved.

6.2 Dynamic Variable Ordering Framework

6.2.1 Watched Variables. To detect univariate clauses and lemmas, we implement two watched variables inspired by two watched literals. Each clause or lemma is watched by two variables (boolean or arithmetic) appearing in it. Watchers are changed when one of them is being assigned. Cases are split below:

- there exists a third variable unassigned: we just replace the assigned watcher by this one.
- there is no other variable unassigned: this clause must be univariate to the unassigned watcher.
- both watchers are assigned: we do nothing for this case.

Whenever we detect a univariate clause, the feasible set is updated in an eager way, and thus helps the search engine gather more clause-level information.

6.2.2 Projection Order. A usual problem faced by nonlinear arithmetic is the strict variable order relationship of root atoms, which are generated by model-based projection. Given a polynomial set ps, and a projection order $\{v_1, v_2, ..., v_k\}$, each time the projection method eliminates a variable and generates a root atom according to that variable.

As discussed in [43], in most cases the variable outside the polynomial needs to be assigned at last. Specifically, when all atoms are in root format, the projection order should exactly be the inverse order of the assignment, as implemented in our solver.

6.2.3 Branching Heuristic. VSIDS is a particularly effective branching heuristic. We increase activities of variables each time a conflict is detected. All variables involved in the conflict analysis will be

[6] For look-ahead mechanism, this happens only for block cases.

Symbol	Description	Value
arith_decay	Decay_factor for arithmetic variables in VSIDS	0.95
bool_decay	Decay_factor for boolean variables in VSIDS	0.95
arith_bump	Incremental amount of arithmetic activity	1
bool_bump	Incremental amount of boolean activity	1
lemma_conf	Initial conflict times for deleting lemmas	100
lemma_conf_inc	Incremental factor of conflict for lemmas	1.5

Table 1: Tunable parameters

considered. We design several branching heuristics as suggested by [43].

- **Boolean First**: arithmetic variables are only assigned after all boolean variables are assigned. In fact, original static order NLSAT always branch boolean variable at first.
- **Arithmetic First**: boolean variables are only assigned after all arithmetic variables are assigned.
- **Uniform**: Whatever type the variable is, they are ordered only by their activity.

By comparing experimental results, We use uniform heuristic in our implementation.

6.2.4 Lemma Management. We record the activity of learned lemmas in conflict analysis, and periodically delete half of them. Compared with sat solvers, lemmas are not usually all useful due to the existence of root atoms. Deleting lemmas actually decreases memory allocation spent on useless lemmas.

6.2.5 Parameter Settings. Values of tunable parameters are summarized in the Table 1.

6.3 Shortcut for UNSAT Instances

For the block case in Section 4, we process blocked clauses as the same with NLSAT. As concluded previously, this happens because previous variables (arithmetic or boolean) takes a wrong value. In the real implementation, a shortcut mechanism is added to directly return unsat if the blocked clauses only contain one variable. In this case, there is no previous stages and thus the instance must be unsatisfiable.

7 EVALUATION

In this section, we compare our algorithm with other existing solvers, such as Z3 (version 4.13.1) [20], CVC5 (version 1.0.2) [6] and YICES2 (version 2.6.2) [22]. Ablation study of several improvements are also described.

7.1 Experiment Preliminaries

The normally used benchmark for evaluating SMT solvers is SMT-LIB. The whole benchmark of QF_NRA theory consists of 12134 instances, from various applications ranging from nonlinear hybrid automata, ranking functions generating for program analysis, and also mathematical problems. Most instances have been labelled as satisfiable or unsatisfiable, but still some are labelled as unknown. One should note that instances from SMT-LIB always show different forms in clause numbers, literal numbers and even polynomial degrees. Our experiments are evaluated on a server with Intel Xeon Platinum 8153 processor at 2.00 GHz. We limit the running time of each instance 1200 seconds, just the same in the SMT-COMP.

Category	#inst		Z3	YICES2	CVC5	NLSAT	Ours
20161105-Sturm-MBO	405	SAT	0	0	0	0	0
		UNSAT	124	**285**	**285**	44	39
		SOLVED	124	**285**	**285**	44	39
20161105-Sturm-MGC	9	SAT	**2**	0	0	**2**	**2**
		UNSAT	**7**	0	0	**7**	6
		SOLVED	**9**	0	0	**9**	8
20170501-Heizmann	69	SAT	**2**	0	1	1	**2**
		UNSAT	1	12	9	10	**19**
		SOLVED	3	12	10	11	**21**
20180501-Economics-Mulligan	135	SAT	**93**	91	89	**93**	92
		UNSAT	39	39	35	**41**	**41**
		SOLVED	132	130	124	**134**	131
2019-ezsmt	63	SAT	56	52	50	**58**	36
		UNSAT	**2**	**2**	**2**	**2**	**2**
		SOLVED	58	54	52	**60**	38
20200911-Pine	245	SAT	234	**235**	199	**235**	234
		UNSAT	6	**8**	5	7	5
		SOLVED	240	**243**	204	242	239
20211101-Geogebra	112	SAT	**110**	99	91	**110**	98
		UNSAT	0	0	0	0	0
		SOLVED	**110**	99	91	**110**	98
20220314-Uncu	225	SAT	69	**70**	62	68	**70**
		UNSAT	**155**	153	148	**155**	152
		SOLVED	**224**	223	210	223	222
hong	20	SAT	0	0	0	0	0
		UNSAT	8	**20**	**20**	12	14
		SOLVED	8	**20**	**20**	12	14
hycomp	2752	SAT	**307**	227	225	244	291
		UNSAT	**2242**	2201	2212	2088	2181
		SOLVED	**2549**	2428	2437	2332	2472
kissing	45	SAT	**33**	10	17	12	14
		UNSAT	0	0	0	0	0
		SOLVED	**33**	10	17	12	14
LassoRanker	821	SAT	167	122	**305**	220	302
		UNSAT	151	260	**470**	174	311
		SOLVED	318	382	**775**	394	613
meti-tarski	7006	SAT	**4391**	4369	4343	**4391**	4372
		UNSAT	2605	2588	2581	**2611**	2588
		SOLVED	6996	6957	6924	**7002**	6960
UltimateAutomizer	61	SAT	35	39	35	**45**	39
		UNSAT	11	12	10	**13**	12
		SOLVED	46	51	45	**58**	51
zankl	166	SAT	**70**	58	58	62	56
		UNSAT	28	**32**	**32**	27	30
		SOLVED	**98**	90	90	89	86
Total	12134	SAT	5569	5372	5475	5541	**5608**
		UNSAT	5379	5612	**5809**	5191	5397
		SOLVED	10948	10984	**11284**	10732	11005

Table 2: Summary of results for all instances in SMT-LIB (QF_NRA).

7.2 Overall Result

We compare our algorithm with other SMT solvers in Table 2. Note that when evaluating our algorithm in Z3 solver, we disable all other tactics like CDCL(T) and incomplete algorithms. Solvers including Z3, CVC5 and YICES2 are all tested without any modification, carrying a portfolio of different algorithms. We also test original NLSAT solver by disabling other tactics.

Our algorithm is competitive with current state-of-the-art solvers, like Z3 and CVC5. Specifically, we solve the most satisfiable instances and the second most unsatisfiable instances. Scatter plots about solution times are shown in Figure 5.

7.2.1 Comparison with CVC5. CVC5 solves the most instances in our experiment. However, it is especially effective in unsat instances thanks to incomplete algorithms like interval constraint propagation and incremental linearization. A traditional category, called MBO [3], contains only one clause with a very high degree, which makes it difficult to be solved by CAD-based algorithms.

7.2.2 Comparison with Original NLSAT. Actually, our solver decreases a little in most instances. These instances always contain

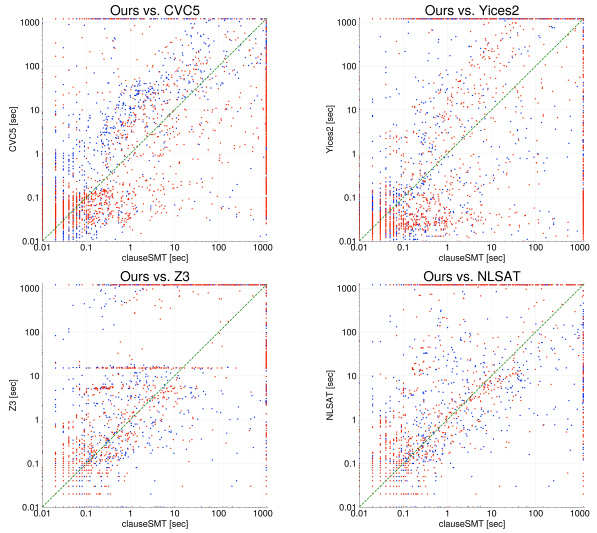

Figure 5: Run time comparison against CVC5, Z3, YICES2 and nlsat (blue points: satisfiable instances, red points: unsatisfiable instances).

high degree polynomials and thus make our feasible set computation relatively heavy. However, our solver shows a strong advancement in LassoRanker and Hycomp. These categories contain thousands of instances and usually have a complex feasible set relationship. Our solver increases a half in LassoRanker, which comes from termination analysis [26, 35]. Our solver has improved NLSAT almost three hundred instances for the overall result.

7.3 Effectiveness of Look-Ahead Mechanism

In order to show the effect of look-ahead mechanism, several versions are implemented, as denoted below. The experimental results are shown in Table 3.

- **Look-Ahead**: Implement feasible set based look-ahead mechanism on original NLSAT solver (static variable order).
- **Lower Degree**: Decide literals with the lowest degree. This is the default heuristic in NLSAT.
- **Random Decide:** Randomly decide literals when processing clauses.

Although the difference of solved problems in most categories is not large, our algorithm solves about 50 more instances in the Hycomp [18] category, which contains a huge amount of nonlinear equalities. In this case, Hycomp instances usually exist in literal path cases, which is precisely the advantage of the look-ahead mechanism.

Theoretically, look-ahead mechanism can detect block cases and avoid conflicts for path cases. We record the conflict times in both algorithms. The scatter plot is shown in Figure 6. The red line represents solving this instance brings the same conflict number for both implementation. As the scatter plot shows, most instances can be solved using fewer times of conflict. Although this property does not bring a huge gap for a 1200 seconds test, it will help boost the systematic search for clauses with multiple literals.

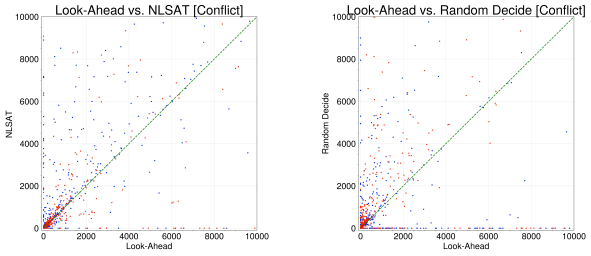

Figure 6: Conflict times of look-ahead NLSAT against original and random NLSAT (blue points: satisfiable instances, red points: unsatisfiable instances).

Category	#inst	Decide Lower Degree	Random Decide	Look-Ahead
20161105-Sturm-MBO	405	44	45	44
20161105-Sturm-MGC	9	9	9	9
20170501-Heizmann	69	11	5	7
20180501-Economics-Mulligan	135	134	134	134
2019-ezsmt	63	60	59	58
20200911-Pine	245	242	242	243
20211101-Geogebra	112	110	109	110
20220314-Uncu	225	223	224	224
hong	20	12	12	12
hycomp	2752	2332	2272	2388
kissing	45	12	14	15
LassoRanker	821	394	393	389
meti-tarski	7006	7002	7001	7002
UltimateAutomizer	61	58	44	57
zankl	166	89	89	87
Total	12134	10732	10652	10778

Table 3: Comparison of solved instances for different literal decision mechanisms.

7.4 Effectiveness of Clause-Level Propagation

Clause-Level Propagation can only be evaluated on dynamic variable ordering framework. Thus, we implement three versions for comparison: original NLSAT solver with static order based on variable's degree (static), dynamic NLSAT solver (VSIDS), and clause-level propagation dynamic NLSAT algorithm (prop-VSIDS). Results are shown in Table 4.

The results indicate that VSIDS is effective in MCSAT framework. Besides that, clause-level propagation indeed speeds up the detection of conflicts, and increases the solved instances in most categories. More concretely, prop-VSIDS is especially competitive in three categories: hycomp [18], LassoRanker [26, 35] and meti-tarski [2]. All these three categories contain a huge number of arithmetic clauses, and thus easily makes the arithmetic variable fall into a block case. Stages (semantic decision times) of prop-VSIDS against traditional VSIDS is shown in Figure 7. By detecting inconsistent branching choices quickly, the overall stages required for a problem has been decreased significantly.

7.5 Threats to Validity

Correctness of implementation. It costs a significant amount of work to build develop our new solver clauseSMT. Comparisons with other competitors are executed on the same environment. The results have been checked in detail just in case our solver outputs an unexpected result. Satisfiable instances are carefully checked by validators to ensure correctness.

Random Characteristics. Although NLSAT uses an internal random tool for semantics decision and reordering for clauses and

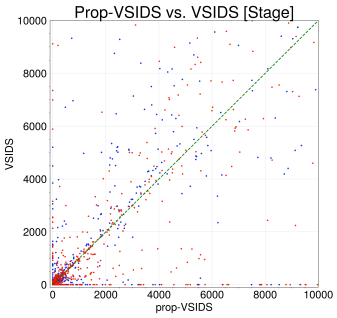

Figure 7: Stage comparison of prop-VSIDS against VSIDS (blue points: satisfiable instances, red points: unsatisfiable instances).

Category	#inst	Static	VSIDS	prop-VSIDS
20161105-Sturm-MBO	405	44	38	39
20161105-Sturm-MGC	9	9	6	8
20170501-Heizmann	69	7	20	21
20180501-Economics-Mulligan	135	134	133	133
2019-ezsmt	63	58	30	38
20200911-Pine	245	243	239	239
20211101-Geogebra	112	110	101	98
20220314-Uncu	225	224	222	222
hong	20	12	11	11
hycomp	2752	2388	2426	2472
kissing	45	15	14	14
LassoRanker	821	389	571	613
meti-tarski	7006	7002	6974	6960
UltimateAutomizer	61	57	52	51
zankl	166	87	83	86
Total	12134	10778	10920	11005

Table 4: Comparison of solved instances for different branching heuristics.

literals, all these implementations are maintained in our solver and thus won't cause any difference. Besides, even for the literal decision part, we give a random version which is not competitive at all as discussed above. Some important parameters like conflict times and stages are also depicted by scatter figures.

8 RELATED WORK

Methods for SMT solving are divided into two main kinds: complete methods and incomplete methods. Incomplete methods are effective due to their high speed and carefully designed techniques. Among them, interval constraint propagation (ICP) [32, 49] is a common method to quickly detect unsatisfiable instances, as implemented in the dReal solver [24]. Recently, local search [27] has also been introduced from SAT problem to arithmetic theories, including integer arithmetic [11, 12], linear and multilinear real arithmetic [36], and nonlinear real arithmetic [39, 51]. Other incomplete methods like incremental linearization [16, 17] and subtropical method [23, 41] have also been explored to solve nonlinear constraints more efficiently. Complete methods are the main part of existing SMT solvers because of their high performance on both satisfiable and unsatisfiable instances. Both CDCL(T) [44] and NLSAT framework [30] use CAD for their theory solver or explanation module respectively [33]. By using other light methods for explanation, MCSAT maintains its performance on more different categories and applications [21]. Until now, this innovative framework is still the main complete method used for different theories, including linear or nonlinear, integer or real arithmetic. Some works trying to improve the effectiveness of NLSAT have also been studied. For example, some works try to decrease the complexity of CAD by learning better projection orders of variables [14, 28, 38], designing innovative projection operators [37], and generating larger cells which are literal-invariant [1, 42]. Other works talk about the dynamic branching heuristic for MCSAT [43]. The proof complexity of MCSAT has also been studied theoretically [34]. Besides, some recent works like hybridSMT [54] have studied the feasibility of incorporating local search into CDCL(T).

9 CONCLUSION

In this paper, we present a NLSAT-based clause-level algorithm for SMT problem over nonlinear real arithmetic. We categorize the conflicts in NLSAT and talk about the literal decision problem that many CDCL-style algorithms would encounter. To deal with the problem, several improvements including feasible set based look-ahead mechanism and clause-level propagation based branching have been presented. Through the experiment, our new algorithm is competitive against many other existing solvers. Besides, we show the effectiveness of proposed mechanisms.

In the future, we hope to find a clause-level method to tackle the block case. Typically, the block case can be reduced to a quantifier elimination problem. There may exist some lighter methods compared with CAD to build the connection with literals in different clauses, and thus change the previous decisions and create a consistent decision path. Inspired by the deep cooperation of local search and CDCL(T) [54], it is also interesting to combine our new algorithm with a clause-level local search algorithm.

REFERENCES

[1] Erika Ábrahám, James H. Davenport, Matthew England, and Gereon Kremer. 2021. Deciding the consistency of non-linear real arithmetic constraints with a conflict driven search using cylindrical algebraic coverings. *J. Log. Algebraic Methods Program.* 119 (2021), 100633. https://doi.org/10.1016/j.jlamp.2020.100633

[2] Behzad Akbarpour and Lawrence C. Paulson. 2010. MetiTarski: An Automatic Theorem Prover for Real-Valued Special Functions. *J. Autom. Reason.* 44, 3 (2010), 175–205. https://doi.org/10.1007/s10817-009-9149-2

[3] Tatsuya Akutsu, Morihiro Hayashida, and Takeyuki Tamura. 2008. Algorithms for Inference, Analysis and Control of Boolean Networks. In *Algebraic Biology, Third International Conference, AB 2008, Castle of Hagenberg, Austria, July 31-August 2, 2008, Proceedings (Lecture Notes in Computer Science, Vol. 5147)*, Katsuhisa Horimoto, Georg Regensburger, Markus Rosenkranz, and Hiroshi Yoshida (Eds.). Springer, 1–15. https://doi.org/10.1007/978-3-540-85101-1_1

[4] Guy Amir, Haoze Wu, Clark Barrett, and Guy Katz. 2021. An SMT-Based Approach for Verifying Binarized Neural Networks. In *Tools and Algorithms for the Construction and Analysis of Systems: 27th International Conference, TACAS 2021, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2021, Luxembourg City, Luxembourg, March 27 – April 1, 2021, Proceedings, Part II* (Luxembourg City, Luxembourg). Springer-Verlag, Berlin, Heidelberg, 203–222. https://doi.org/10.1007/978-3-030-72013-1_11

[5] Kyungmin Bae and Sicun Gao. 2017. Modular SMT-based analysis of nonlinear hybrid systems. In *2017 Formal Methods in Computer Aided Design (FMCAD)*. 180–187. https://doi.org/10.23919/FMCAD.2017.8102258

[6] Haniel Barbosa, Clark W. Barrett, Martin Brain, Gereon Kremer, Hanna Lachnitt, Makai Mann, Abdalrhman Mohamed, Mudathir Mohamed, Aina Niemetz, Andres

Nötzli, Alex Ozdemir, Mathias Preiner, Andrew Reynolds, Ying Sheng, Cesare Tinelli, and Yoni Zohar. 2022. cvc5: A Versatile and Industrial-Strength SMT Solver. In *Tools and Algorithms for the Construction and Analysis of Systems - 28th International Conference, TACAS 2022, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2022, Munich, Germany, April 2-7, 2022, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 13243)*, Dana Fisman and Grigore Rosu (Eds.). Springer, 415–442. https://doi.org/10.1007/978-3-030-99524-9_24

[7] Clark Barrett, Pascal Fontaine, and Cesare Tinelli. 2016. The Satisfiability Modulo Theories Library (SMT-LIB). `www.SMT-LIB.org`.

[8] Clark W. Barrett and Cesare Tinelli. 2018. Satisfiability Modulo Theories. In *Handbook of Model Checking*, Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem (Eds.). Springer, 305–343. https://doi.org/10.1007/978-3-319-10575-8_11

[9] Dirk Beyer, Matthias Dangl, and Philipp Wendler. 2018. A Unifying View on SMT-Based Software Verification. *J. Autom. Reason.* 60, 3 (mar 2018), 299–335. https://doi.org/10.1007/s10817-017-9432-6

[10] Cristian Cadar, Daniel Dunbar, and Dawson Engler. 2008. KLEE: unassisted and automatic generation of high-coverage tests for complex systems programs. In *Proceedings of the 8th USENIX Conference on Operating Systems Design and Implementation* (San Diego, California) *(OSDI'08)*. USENIX Association, USA, 209–224.

[11] Shaowei Cai, Bohan Li, and Xindi Zhang. 2022. Local Search for SMT on Linear Integer Arithmetic. In *Computer Aided Verification - 34th International Conference, CAV 2022, Haifa, Israel, August 7-10, 2022, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 13372)*, Sharon Shoham and Yakir Vizel (Eds.). Springer, 227–248. https://doi.org/10.1007/978-3-031-13188-2_12

[12] Shaowei Cai, Bohan Li, and Xindi Zhang. 2023. Local Search For Satisfiability Modulo Integer Arithmetic Theories. *ACM Trans. Comput. Logic* 24, 4, Article 32 (jul 2023), 26 pages. https://doi.org/10.1145/3597495

[13] B. F. Caviness and Jeremy R. Johnson. 2004. Quantifier Elimination and Cylindrical Algebraic Decomposition. In *Texts and Monographs in Symbolic Computation*.

[14] Changbo Chen, Zhangpeng Zhu, and Haoyu Chi. 2020. Variable Ordering Selection for Cylindrical Algebraic Decomposition with Artificial Neural Networks. In *Mathematical Software – ICMS 2020: 7th International Conference, Braunschweig, Germany, July 13–16, 2020, Proceedings* (Braunschweig, Germany). Springer-Verlag, Berlin, Heidelberg, 281–291. https://doi.org/10.1007/978-3-030-52200-1_28

[15] Alessandro Cimatti. 2012. Application of SMT solvers to hybrid system verification. In *2012 Formal Methods in Computer-Aided Design (FMCAD)*. 4–4.

[16] Alessandro Cimatti, Alberto Griggio, Ahmed Irfan, Marco Roveri, and Roberto Sebastiani. 2018. Experimenting on Solving Nonlinear Integer Arithmetic with Incremental Linearization. In *Theory and Applications of Satisfiability Testing – SAT 2018: 21st International Conference, SAT 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 9–12, 2018, Proceedings* (Oxford, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 383–398. https://doi.org/10.1007/978-3-319-94144-8_23

[17] Alessandro Cimatti, Alberto Griggio, Ahmed Irfan, Marco Roveri, and Roberto Sebastiani. 2018. Incremental Linearization for Satisfiability and Verification Modulo Nonlinear Arithmetic and Transcendental Functions. *ACM Trans. Comput. Log.* 19, 3 (2018), 19:1–19:52. https://doi.org/10.1145/3230639

[18] Alessandro Cimatti, Sergio Mover, and Stefano Tonetta. 2012. A quantifier-free SMT encoding of non-linear hybrid automata. In *Formal Methods in Computer-Aided Design, FMCAD 2012, Cambridge, UK, October 22-25, 2012*, Gianpiero Cabodi and Satnam Singh (Eds.). IEEE, 187–195. https://ieeexplore.ieee.org/document/6462573/

[19] Florian Corzilius, Gereon Kremer, Sebastian Junges, Stefan Schupp, and Erika Ábrahám. 2015. SMT-RAT: An Open Source C++ Toolbox for Strategic and Parallel SMT Solving. In *Theory and Applications of Satisfiability Testing – SAT 2015*, Marijn Heule and Sean Weaver (Eds.). Springer International Publishing, Cham, 360–368.

[20] Leonardo Mendonça de Moura and Nikolaj S. Bjørner. 2008. Z3: An Efficient SMT Solver. In *Tools and Algorithms for the Construction and Analysis of Systems, 14th International Conference, TACAS 2008, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2008, Budapest, Hungary, March 29-April 6, 2008. Proceedings (Lecture Notes in Computer Science, Vol. 4963)*, C. R. Ramakrishnan and Jakob Rehof (Eds.). Springer, 337–340. https://doi.org/10.1007/978-3-540-78800-3_24

[21] Leonardo Mendonça de Moura and Dejan Jovanovic. 2013. A Model-Constructing Satisfiability Calculus. In *Verification, Model Checking, and Abstract Interpretation, 14th International Conference, VMCAI 2013, Rome, Italy, January 20-22, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 7737)*, Roberto Giacobazzi, Josh Berdine, and Isabella Mastroeni (Eds.). Springer, 1–12. https://doi.org/10.1007/978-3-642-35873-9_1

[22] Bruno Dutertre. 2014. Yices 2.2. In *Computer Aided Verification - 26th International Conference, CAV 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 18-22, 2014. Proceedings (Lecture Notes in Computer Science, Vol. 8559)*, Armin Biere and Roderick Bloem (Eds.). Springer, 737–744. https://doi.org/10.1007/978-3-319-08867-9_49

[23] Pascal Fontaine, Mizuhito Ogawa, Thomas Sturm, and Xuan-Tung Vu. 2017. Subtropical Satisfiability. In *Frontiers of Combining Systems - 11th International Symposium, FroCoS 2017, Brasília, Brazil, September 27-29, 2017, Proceedings (Lecture Notes in Computer Science, Vol. 10483)*, Clare Dixon and Marcelo Finger (Eds.). Springer, 189–206. https://doi.org/10.1007/978-3-319-66167-4_11

[24] Sicun Gao, Soonho Kong, and Edmund M. Clarke. 2013. dReal: An SMT Solver for Nonlinear Theories over the Reals. In *Automated Deduction - CADE-24 - 24th International Conference on Automated Deduction, Lake Placid, NY, USA, June 9-14, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 7898)*, Maria Paola Bonacina (Ed.). Springer, 208–214. https://doi.org/10.1007/978-3-642-38574-2_14

[25] Patrice Godefroid, Nils Klarlund, and Koushik Sen. 2005. DART: directed automated random testing. In *Proceedings of the 2005 ACM SIGPLAN Conference on Programming Language Design and Implementation* (Chicago, IL, USA) *(PLDI '05)*. Association for Computing Machinery, New York, NY, USA, 213–223. https://doi.org/10.1145/1065010.1065036

[26] Matthias Heizmann, Jochen Hoenicke, Jan Leike, and Andreas Podelski. 2013. Linear Ranking for Linear Lasso Programs. In *Automated Technology for Verification and Analysis - 11th International Symposium, ATVA 2013, Hanoi, Vietnam, October 15-18, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 8172)*, Dang Van Hung and Mizuhito Ogawa (Eds.). Springer, 365–380. https://doi.org/10.1007/978-3-319-02444-8_26

[27] Holger H. Hoos and Thomas Stützle. 2004. *Stochastic Local Search: Foundations & Applications*. Elsevier / Morgan Kaufmann.

[28] Fuqi Jia, Yuhang Dong, Minghao Liu, Pei Huang, Feifei Ma, and Jian Zhang. 2023. Suggesting Variable Order for Cylindrical Algebraic Decomposition via Reinforcement Learning. In *Thirty-seventh Conference on Neural Information Processing Systems*. https://openreview.net/forum?id=vNsdFwjPtL

[29] Dejan Jovanovic, Clark Barrett, and Leonardo de Moura. 2013. The design and implementation of the model constructing satisfiability calculus. In *2013 Formal Methods in Computer-Aided Design*. 173–180. https://doi.org/10.1109/FMCAD.2013.7027033

[30] Dejan Jovanovic and Leonardo Mendonça de Moura. 2012. Solving Non-linear Arithmetic. In *Automated Reasoning - 6th International Joint Conference, IJCAR 2012, Manchester, UK, June 26-29, 2012. Proceedings (Lecture Notes in Computer Science, Vol. 7364)*, Bernhard Gramlich, Dale Miller, and Uli Sattler (Eds.). Springer, 339–354. https://doi.org/10.1007/978-3-642-31365-3_27

[31] Guy Katz, Clark Barrett, David L. Dill, Kyle Julian, and Mykel J. Kochenderfer. 2017. Reluplex: An Efficient SMT Solver for Verifying Deep Neural Networks. In *Computer Aided Verification*, Rupak Majumdar and Viktor Kunčak (Eds.). Springer International Publishing, Cham, 97–117.

[32] To Van Khanh and Mizuhito Ogawa. 2012. SMT for Polynomial Constraints on Real Numbers. In *Third Workshop on Tools for Automatic Program Analysis, TAPAS 2012, Deauville, France, September 14, 2012 (Electronic Notes in Theoretical Computer Science, Vol. 289)*, Bertrand Jeannet (Ed.). Elsevier, 27–40. https://doi.org/10.1016/j.entcs.2012.11.004

[33] Gereon Kremer. 2020. *Cylindrical algebraic decomposition for nonlinear arithmetic problems*. Ph. D. Dissertation. RWTH Aachen University, Germany. https://publications.rwth-aachen.de/record/792185

[34] Gereon Kremer, Erika Abraham, and Vijay Ganesh. 2021. On the proof complexity of MCSAT. arXiv:2109.01585 [cs.LO]

[35] Jan Leike and Matthias Heizmann. 2015. Ranking Templates for Linear Loops. *Log. Methods Comput. Sci.* 11, 1 (2015). https://doi.org/10.2168/LMCS-11(1:16)2015

[36] Bohan Li and Shaowei Cai. 2023. Local Search For SMT On Linear and Multi-linear Real Arithmetic. In *2023 Formal Methods in Computer-Aided Design (FMCAD)*. 1–10. https://doi.org/10.34727/2023/isbn.978-3-85448-060-0_25

[37] Haokun Li and Bican Xia. 2020. Solving Satisfiability of Polynomial Formulas By Sample-Cell Projection. arXiv:2003.00409 [cs.LO]

[38] Haokun Li, Bican Xia, Huiying Zhang, and Tao Zheng. 2023. Choosing better variable orderings for cylindrical algebraic decomposition via exploiting chordal structure. *J. Symb. Comput.* 116 (2023), 324–344. https://doi.org/10.1016/j.jsc.2022.10.009

[39] Haokun Li, Bican Xia, and Tianqi Zhao. 2023. Local Search for Solving Satisfiability of Polynomial Formulas. In *Computer Aided Verification - 35th International Conference, CAV 2023, Paris, France, July 17-22, 2023, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 13965)*, Constantin Enea and Akash Lal (Eds.). Springer, 87–109. https://doi.org/10.1007/978-3-031-37703-7_5

[40] M.W. Moskewicz, C.F. Madigan, Y. Zhao, L. Zhang, and S. Malik. 2001. Chaff: engineering an efficient SAT solver. In *Proceedings of the 38th Design Automation Conference (IEEE Cat. No.01CH37232)*. 530–535. https://doi.org/10.1145/378239.379017

[41] Jasper Nalbach and Erika Ábrahám. 2023. Subtropical Satisfiability for SMT Solving. In *NASA Formal Methods - 15th International Symposium, NFM 2023, Houston, TX, USA, May 16-18, 2023, Proceedings (Lecture Notes in Computer Science, Vol. 13903)*, Kristin Yvonne Rozier and Swarat Chaudhuri (Eds.). Springer, 430–446. https://doi.org/10.1007/978-3-031-33170-1_26

[42] Jasper Nalbach, Erika Ábrahám, Philippe Specht, Christopher W. Brown, James H. Davenport, and Matthew England. 2024. Levelwise construction of a single cylindrical algebraic cell. *J. Symb. Comput.* 123, C (may 2024), 44 pages. https://doi.org/10.1016/j.jsc.2023.102288

[43] Jasper Nalbach, Gereon Kremer, and Erika Ábrahám. 2019. On Variable Orderings in MCSAT for Non-Linear Real Arithmetic. In *SC-square@SIAM AG*. https://api.semanticscholar.org/CorpusID:204767299

[44] Robert Nieuwenhuis, Albert Oliveras, and Cesare Tinelli. 2006. Solving SAT and SAT Modulo Theories: From an abstract Davis–Putnam–Logemann–Loveland procedure to DPLL(T). *J. ACM* 53, 6 (2006), 937–977. https://doi.org/10.1145/1217856.1217859

[45] Brandon Paulsen and Chao Wang. 2022. Example Guided Synthesis of Linear Approximations for Neural Network Verification. In *Computer Aided Verification: 34th International Conference, CAV 2022, Haifa, Israel, August 7–10, 2022, Proceedings, Part I* (Haifa, Israel). Springer-Verlag, Berlin, Heidelberg, 149–170. https://doi.org/10.1007/978-3-031-13185-1_8

[46] Brandon Paulsen and Chao Wang. 2022. LinSyn: Synthesizing Tight Linear Bounds for Arbitrary Neural Network Activation Functions. In *Tools and Algorithms for the Construction and Analysis of Systems: 28th International Conference, TACAS 2022, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2022, Munich, Germany, April 2–7, 2022, Proceedings, Part I* (Munich, Germany). Springer-Verlag, Berlin, Heidelberg, 357–376. https://doi.org/10.1007/978-3-030-99524-9_19

[47] Yasser Shoukry, Michelle Chong, Masashi Wakaiki, Pierluigi Nuzzo, Alberto Sangiovanni-Vincentelli, Sanjit A. Seshia, JOÃo P. Hespanha, and Paulo Tabuada. 2018. SMT-Based Observer Design for Cyber-Physical Systems under Sensor Attacks. *ACM Trans. Cyber-Phys. Syst.* 2, 1, Article 5 (jan 2018), 27 pages. https://doi.org/10.1145/3078621

[48] Martin Tappler, Bernhard K. Aichernig, and Florian Lorber. 2022. Timed Automata Learning via SMT Solving. In *NASA Formal Methods: 14th International Symposium, NFM 2022, Pasadena, CA, USA, May 24–27, 2022, Proceedings* (Pasadena, CA, USA). Springer-Verlag, Berlin, Heidelberg, 489–507. https://doi.org/10.1007/978-3-031-06773-0_26

[49] Vu Xuan Tung, To Van Khanh, and Mizuhito Ogawa. 2017. raSAT: an SMT solver for polynomial constraints. *Formal Methods Syst. Des.* 51, 3 (2017), 462–499. https://doi.org/10.1007/s10703-017-0284-9

[50] Jingbo Wang and Chao Wang. 2023. Learning to Synthesize Relational Invariants. In *Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering* (Rochester, MI, USA) *(ASE '22)*. Association for Computing Machinery, New York, NY, USA, Article 65, 12 pages. https://doi.org/10.1145/3551349.3556942

[51] Zhonghan Wang, Bohua Zhan, Bohan Li, and Shaowei Cai. 2024. Efficient Local Search for Nonlinear Real Arithmetic. In *Verification, Model Checking, and Abstract Interpretation: 25th International Conference, VMCAI 2024, London, United Kingdom, January 15–16, 2024, Proceedings, Part I* (London, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 326–349. https://doi.org/10.1007/978-3-031-50524-9_15

[52] Runqing Xu, Jie An, and Bohua Zhan. 2022. Active Learning of One-Clock Timed Automata Using Constraint Solving. In *Automated Technology for Verification and Analysis: 20th International Symposium, ATVA 2022, Virtual Event, October 25–28, 2022, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 249–265. https://doi.org/10.1007/978-3-031-19992-9_16

[53] Peisen Yao, Qingkai Shi, Heqing Huang, and Charles Zhang. 2021. Program analysis via efficient symbolic abstraction. *Proc. ACM Program. Lang.* 5, OOPSLA, Article 118 (oct 2021), 32 pages. https://doi.org/10.1145/3485495

[54] Xindi Zhang, Bohan Li, and Shaowei Cai. 2024. Deep Combination of CDCL(T) and Local Search for Satisfiability Modulo Non-Linear Integer Arithmetic Theory. In *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering* (Lisbon, Portugal) *(ICSE '24)*. Association for Computing Machinery, New York, NY, USA, Article 125, 13 pages. https://doi.org/10.1145/3597503.3639105